Research Note 82-30

AD A130246

DEVELOPMENT OF A METHODOLOGY FOR CONDUCTING TRAINING EFFECTIVENESS EVALUATIONS OF AIR DEFENSE TRAINING, AND ABSTRACTS OF TEE-RELATED LITERATURE

Robert P. Fishburne, Jr., Steven J. Rolnick
and Joseph Y. Larsen, III
Calspan Corporation
Advanced Technology Center

ARI FIELD UNIT AT FORT BLISS, TEXAS

U. S. Army
Research Institute for the Behavioral and Social Sciences

June 1982

Unclassified

SECURITY CLASSIFICATION OF THIS PAGE (When Data Entered)

REPORT DOCUMENTATION PAGE

READ INSTRUCTIONS BEFORE COMPLETING FORM

1. REPORT NUMBER: Research Note 82-30	2. GOVT ACCESSION NO.: AD-A130246
3. RECIPIENT'S CATALOG NUMBER:	
4. TITLE (and Subtitle): DEVELOPMENT OF A METHODOLOGY FOR CONDUCTING TRAINING EFFECTIVENESS EVALUATIONS OF AIR DEFENSE TRAINING, AND ABSTRACTS OF TEE-RELATED LITERATURE	5. TYPE OF REPORT & PERIOD COVERED: Final Report
	6. PERFORMING ORG. REPORT NUMBER: --
7. AUTHOR(s): Robert P. Fishburne, Jr. Steven J. Rolnick Joseph Y. Larsen, III	8. CONTRACT OR GRANT NUMBER(s): MDA903-80-C-0579
9. PERFORMING ORGANIZATION NAME AND ADDRESS: Calspan Corporation Advanced Technology Center P.O. Box 400 Buffalo, N.Y. 14225	10. PROGRAM ELEMENT, PROJECT, TASK AREA & WORK UNIT NUMBERS: 2Q263739A793
11. CONTROLLING OFFICE NAME AND ADDRESS: Commandant U. S. Army Air Defense School ATTN: ATSA-TDI, Ft Bliss, TX 79916	12. REPORT DATE: June 1982 13. NUMBER OF PAGES: 103
14. MONITORING AGENCY NAME & ADDRESS (if different from Controlling Office): U.S. Army Research Institute for the Behavioral and Social Sciences (PERI-SB) 5001 Eisenhower Ave., Alexandria, VA. 22333	15. SECURITY CLASS. (of this report): Unclassified 15a. DECLASSIFICATION/DOWNGRADING SCHEDULE: --

16. DISTRIBUTION STATEMENT (of this Report)

Approved for public release, distribution unlimited

17. DISTRIBUTION STATEMENT (of the abstract entered in Block 20, if different from Report)

--

18. SUPPLEMENTARY NOTES

This project was monitored technically by Richard Carter and Gary Sarli, US ARI Field Unit, Fort Bliss, TX.

19. KEY WORDS (Continue on reverse side if necessary and identify by block number)

Training effectiveness evaluation	Product evaluation
Training effectiveness	Air defense training
Training evaluation	Operational tests
Process evaluation	Instructional systems development

20. ABSTRACT (Continue on reverse side if necessary and identify by block number)

The development of a system for conducting training effectiveness evaluations (TEE) on Army Air Defense training is described in detail. Background of the research and development requirement is presented together with a discussion of the purpose, scope, and context of the TEE. A series of events related to the TEE development process is identified, and the literature base is examined in depth. Guidelines for conducting the TEE are outlined by phases and steps, and a master list of evaluation questions is included. The report concludes

DD FORM 1 JAN 73 1473 EDITION OF 1 NOV 65 IS OBSOLETE

Unclassified

SECURITY CLASSIFICATION OF THIS PAGE (When Data Entered)

20. ABSTRACT

with a discussion of the TEE development effort and recommendations for validation and follow-on research and development requirements.

The TEE system described in this report incorporates the principles of instructional systems development and provides for both product evaluation and process evaluation components. Documentation in support of the TEE system includes a proceduralized user's guide entitled "Guidelines for Conducting a Training Effectiveness Evaluation." The user's guide is in two volumes: Volume I is the "TEE Evaluator's Handbook" and Volume II is the "Data Collector's Manual." All materials necessary for the conduct of a TEE, including training materials, job performance aids, and work sheets, are included in the user's guide.

Research Note 82-30

DEVELOPMENT OF A METHODOLOGY FOR CONDUCTING TRAINING EFFECTIVENESS EVALUATIONS OF AIR DEFENSE TRAINING, AND ABSTRACTS OF TEE-RELATED LITERATURE

Robert P. Fishburne, Jr., Steven J. Rolnick
and Joseph Y. Larsen, III
Calspan Corporation
Advanced Technology Center

Submitted by:
Michael H. Strub, Chief
ARI FIELD UNIT AT FORT BLISS, TEXAS

Approved by:
Edgar M. Johnson, Director
SYSTEMS RESEARCH LABORATORY

U.S. ARMY RESEARCH INSTITUTE FOR THE BEHAVIORAL AND SOCIAL SCIENCES
5001 Eisenhower Avenue, Alexandria, Virginia 22333

Office, Deputy Chief of Staff for Personnel
Department of the Army

June 1982

Army Project Number
2Q263739A793

Systems Training

DEVELOPMENT OF A METHODOLOGY FOR CONDUCTING TRAINING EFFECTIVENESS EVALUATIONS OF AIR DEFENSE TRAINING, AND ABSTRACTS OF TEE-RELATED LITERATURE

TABLE OF CONTENTS

SECTION PAGE

LIST OF TABLES

LIST OF FIGURES

SECTION I

Introduction

Background

The Army is in the process of developing and fielding four new Air Defense weapon systems. These are STINGER, ROLAND, SGT York, and PATRIOT. Each of these weapon systems has the requirement for operator and maintenance training programs designed to transition soldiers to the new equipment and to train newly accessioned soldiers. These training programs are to be implemented concurrently with the fielding of each system.

Three major needs have been expressed by the Air Defense community concerning the development of these training programs. They are:

a. A methodology for conducting training effectiveness evaluations (TEE) of Air Defense training packages during operational tests (OT) of the weapon system.

b. Materials and methodologies for performing training effectiveness testing (TET) of Air Defense training devices during the validation, verification, and system integration phases.

c. Methodologies for applying feedback information from Air Defense TEEs and TET to improve training packages and devices.

The Army Research Institute (ARI) has undertaken a program to address each of these needs. The means to fulfill the need for a TEE methodology has been the subject of research performed by Calspan Corporation's Advanced Technology Center under the technical supervision of the Army Research Institute Field Unit at Fort Bliss. Calspan's TEE research is documented in this report.

Investigations by ARI into the need for TEE developments determined that existing methodologies and techniques for evaluating the effectiveness of training packages during OTs were not well articulated. The Army regulations, manuals, and handbooks on TEEs were found to be explicit in what needs to be done, but not in how to do it. Since the typical TEE analyst is not a sophisticated educational technologist, this lack of detailed guidance was determined to be a threat to the accuracy of the TEEs that will be conducted on new training packages under development.

Prior to the TEE research an initial effort to address this need was undertaken by the Army Research Institute Field Unit at Fort Knox (ARIFUFK). That research resulted in the development of prototype TEE materials using the training package developed for OT II of the XM1 main battle tank as a test bed (Harless Performance Guild, Note 1). Guidelines for the following were detailed:

a. Planning the TEE.
b. Observing training and testing.
c. Assessing the quality of trainee performance.
d. Hypothesizing and investigating causes of deficiencies.
e. Documenting the TEE.

The "Harless Guidelines" are discussed in depth in section III of the present report.

Purpose and Scope

The purpose of the TEE research was to refine and validate the Harless methodology for application to Army Air Defense. In approaching this task, the researchers first applied the Harless Guidelines (from the ARIFUFK study) to a sample of OT II training packages apart from their implementation. However, it was soon evident that while the Harless Guidelines were well developed for a process evaluation (i.e., analysis of the conduct of training), they were not applicable to a product evaluation (i.e., analysis of training materials). Since the need remained for a TEE methodology to function in a certification mode when applied to OT training packages prior to their implementation, efforts were broadened to incorporate the development of product evaluation components to the overall TEE procedures. This approach had the additional advantage of laying the groundwork for future development efforts in which the need for tracing performance discrepancies discovered by TEEs back to their causes could be satisfied. Thus, the addition of the product TEE components was seen to have great potential for facilitating the revision process.

The TEE system should be applicable to Army training in general, although it was specifically developed for Army Air Defense OT training. This generalization is based upon its use of ongoing Air Defense training as well as OT training as a test bed during its development, and upon its use of a broad base of evaluation literature for its concepts and procedures. One aspect of Air Defense training which the TEE system addresses is the requirement for evaluating team (i.e., collective) training. Since all Air Defense training is not team training, however, the system does not limit itself to that domain. Most of the procedures for evaluating individual and collective training have more similarities than differences, and the system is adaptable to the unique requirements of most Army training situations. Flexibility is provided in the system through its provisions for tailoring the TEE in accordance with varying requirements and purposes.

Context

As indicated in the discussion of purpose and scope, the TEE system has been developed for applicability to Army Air Defense training and training for operational tests. Additionally, it appears that the TEE system fulfills many of the objectives of the broader scheme of the Training and Doctrine Command (TRADOC) training effectiveness analysis (TEA) system. Descriptions of Air Defense training, OT training, and the TRADOC TEA system therefore follow to provide the reader an understanding of the context in which the TEE system is to be applied.

Air Defense training is driven by Department of Defense directives and Army regulations. The keystone of current training and training development activities is the instructional systems development (ISD) process articulated in TRADOC Pamphlet 350-30, Interservice Procedures for Instructional Systems Development (IPISD) (Branson, Rayner, Cox, Forman, King, and Hannum, 1975). The essential features of the ISD concept as they relate to Air Defense training have been described by Wessling (1979), as follows:

a. It is a total process that includes all elements of a training system.

b. It is focused on the identification of critical tasks for which individual soldiers are to be trained.

c. It takes into account the stratification of the five enlisted skill levels to train only for those tasks immediately applicable.

d. It is focused on the design of each course and lesson based on a validated task listing of the job requirements for each military occupational specialty (MOS).

e. It encourages analysis of competing media to enable the student to gain maximum benefit of time, subject matter, and media selection.

f. It delineates responsibilities for the five separate phases of training developments: task analysis, design, development, implementation (instruction), and evaluation.

g. It provides common linkage to all supporting literature efforts through the constant focus of attention on the critical tasks selected for training.

h. It encourages feedback from all sources to adopt or change the task ordering and media selection.

i. It integrates the feedback from skill qualification tests (SQT) as a measure of training effectiveness of the courses and supporting materials at the school and in the field.

j. It provides a clear linkage between individual tasks and collective tasks for training and evaluation.

k. It provides a context for evaluation of the need for trainees to have training devices and simulators as part of the media selection process.

l. It expands the methods of instruction from the old standard lecture, conference, demonstration, and practical exercise to a greater selection of other media options.

Each enlisted MOS and officer specialty code is represented by a training program based on task analysis. It is supported at specified skill levels by a program of instruction (POI), supporting literature (written at a reading level compatible to the Army grade structure), job performance aids, skill performance aids, and a training strategy. Completion of a POI results in the awarding of the appropriate MOS to enlisted personnel or specialty code to officers.

Supporting training literature may include: soldier's manual, commander's manual, job book, and the Army training and evaluation program (ARTEP). These materials provide a complete listing of job tasks that must be mastered by each soldier. The SQT evaluates the soldier's mastery of these performance tasks. The employment and doctrine guidance for a specific weapon system are provided in field manuals for operational training. Training circulars provide guidance for training personnel on the use of specialized equipment or techniques of employment. ARTEP provides the criteria for measuring unit proficiency in the field as the result of the prescribed collective training program. All of these supporting materials are coordinated in purpose and scope to support individual and collective training.

The TEE methodology has been developed to be complementary to the ISD model and TRADOC Pamphlet 350-30. Additionally, Air Defense POIs, supporting literature, job aids, and training strategy have been addressed in the evaluation questions and in the TEE procedures. Thus applicability to Air Defense training has been insured. Furthermore, a formative evaluation at Fort Bliss has provided verification.

The TSM Guide to Training Development and Acquisition for Major Systems (Hanson and Purifoy, 1978) describes training for operational tests as follows:

a. OT I should evaluate the feasibility of the "training concept" as it is described in the outline individual/collective training plan. OT I should also include studies to insure that training objectives, to which training materials are geared, are valid, i.e., that individuals trained to the objectives can perform at a level consistent with system needs. Other OT I training concerns may include: studies to generate data for cost and training effectiveness analysis (CTEA)/cost and operational effectiveness analysis (COEA), and evaluation of plans for continued training development.

b. OT II will evaluate the capability of the total training package to provide the required training (validation of training objectives) and assess the feasibility of implementing the proposed individual/collective training program.

The TEE methodology has been designed for applicability to OTs in two ways. First, the provision for a product evaluation allows the certification of a training package as to its adequacy for OT training. Second, results of the overall TEE can be used to factor out training-related discrepancies when analyzing OT data.

TRADOC Regulation 350-4, The TRADOC TEA System (Department of Army, 1979a), and the TRADOC TEA Handbook (Department of Army, 1979b) provide the basis for a comprehensive Army training evaluation system, oriented around the life cycle systems management model (LCSMM). Ten objectives have been identified for the TEA System:

a. Assist in the optimization of the soldier-machine and soldier-training subsystem interfaces to enhance battlefield effectiveness.

b. Increase the effectiveness of the training subsystems developments process.

c. Increase combat developments/training developments interface early in and throughout the acquisition process.

d. Increase assurance that the analysis, design, and development phases of ISD are accomplished in a timely manner — before system fielding.

e. Provide, within resource limits, readily accessible analytical assistance to TRADOC schools/agencies engaged in TEA work.

f. Improve resolution of COEA through inclusion of more precise/useful CTEA input.

g. Provide baseline data on generally similar systems for inclusion in considerations of developing systems.

h. Minimize duplication of effort and/or redundancy of resource expenditure.

i. Develop a useful TEA data base.

j. Provide for the organization and coordination of the TEA efforts of TRADOC schools/agencies.

Five different types of TEAs have been developed:

a. Cost and training effectiveness analysis.

b. Initial screening training effectiveness analysis (ISTEA).

c. Training subsystem effectiveness analysis (TSEA).

d. Training developments study.

e. Total system evaluation.

The TEE Guidelines appear to be highly complementary to two types of TEAs, the ISTEA and the TSEA. The purposes of these TEAs are summarized below.

Initial screening training effectiveness analyses are conducted after a system has been fielded in order to:

a. Determine if actual effectiveness and design effectiveness are essentially equal or if a significant performance gap exists.

b. Determine if a cause and effect relationship exists between demonstrated soldier proficiency and attitudes and trainer proficiency and attitudes.

c. Examine aspects of the training environment which are most likely related to the actual/design effectiveness relationship.

Training subsystems effectiveness analyses are also conducted after a system has been fielded in order to:

a. Examine the training subsystem in detail.

b. Determine if the existing significant performance gap is caused, totally or in part, by the training subsystem.

c. Relate soldier, trainer, training environment, training subsystem, and hardware subsystem factors/variables to obtain a high resolution of problem areas.

d. Examine, by excursion, related subsystems (personnel and logistical support subsystems) that may be contributing agents to a performance gap.

e. Identify potential solutions to training subsystems problems.

SECTION II

Development of the TEE Methodology

Approach to TEE Methodology Development

Development of the training effectiveness evaluation methodology was approached from the perspective of both conceptual and procedural considerations. Conceptual considerations included the definition of TEE purposes and the identification of TEE requirements. Procedural considerations involved the design, development, evaluation, and revision of the TEE system. Each of these considerations is discussed in the following paragraphs.

Conceptual Considerations. The overall purposes for the TEE methodology were established as follows:

a. To validate and/or document deficiencies in training for operational tests.

b. To provide an evaluation methodology adapted to the specific requirements of the Air Defense community.

c. To provide an input to a methodology for tracing deficient results identified during Air Defense TEEs back to training system causes.

d. To provide an input to a methodology for improving existing Air Defense courseware generation procedures.

e. To build upon development efforts of Army Research Institute Field Unit - Ft. Knox in the evaluation area.

These purposes and the requirement to build upon a baseline established by the Harless Guidelines provided general direction for the TEE development effort.

In order to achieve the TEE purposes, several components were identified as crucial requirements for the TEE methodology. The first requirement was that the TEE methodology must incorporate product evaluation as well as process evaluation components. As discussed in detail in Secion III, the Harless Guidelines were found to be essentially adequate as a baseline for the process evaluation component. However, product evaluation was not within their scope. Since a well established product evaluation methodology, the "Instructional Quality Inventory" (IQI) (Merrill, Reigeluth, and Faust, 1979) appeared to have potential for fulfilling this requirement, the IQI was selected as the baseline for the product evaluation component.

A second requirement which was identified for the TEE methodology was that team training as well as individual training was a necessary consideration in the Air Defense and OT settings. Examination of the Harless Guidelines revealed that while individual training was evaluated in a rather thorough manner, treatment of team training was not a major consideration.

It was therefore necessary to search the literature for suitable methodological components to augment the TEE baseline in the area of team training. Section III discusses two sources, Thurmond (1980) and Wagner, Nibbits, Rosenblat, and Schultz (1977), which provided some basic guidance in this area.

The third requirement identified for inclusion in the TEE methodology was a thorough treatment of testing issues. The Harless Guidelines were found to be adequate as a baseline for coverage of testing issues in a process evaluation, but not in a product evaluation. Accordingly, the literature was searched for guidance in this area. Section III identifies four sources that contributed to the TEE methodology in evaluating the adequacy of test items and test administration: Swezey and Pearlstein (1975); Roid and Haladyna (Note 3); Courseware (Note 4); and U. S. Navy (1976).

The fourth and final major requirement which was identified conceptually for consideration in development of the TEE methodology was the incorporation of an ISD model as a frame of reference. This approach was adopted to insure the inclusion of principles of instructional technology and to facilitate the revision process as a follow-on activity to the TEE methodology. The discussion in Section III identifies The Interservice Procedures for Instructional Systems Development (Branson, et al. 1975) as the model selected.

Procedural Considerations. The design of the TEE procedures followed the general structure of the Harless Guidelines, but with a few noteworthy differences. The first and most important was the decision to develop a master list of evaluation questions from which the TEE analyst could select for either product or process evaluation. This procedure was incorporated in the TEE methodology to allow tailoring to the specific needs of each evaluation purpose and context.

The second difference from the Harless Guidelines in the TEE design was to replace the procedure for hypothesizing causes of performance discrepancies (i.e., the Harless procedure for testing the skill/knowledge hypothesis) with a more direct approach. An alternate procedure was designed to identify performance discrepancies from test data and training deficiencies from comparisons of training program components with standards. The goal of this change of procedures was to facilitate the revision process by tracing performance discrepancies to their probable causes (e.g., specific training deficiencies).

Each of the previously discussed conceptual requirements for TEE methodology development was also incorporated into the procedural design. The addition of the product evaluation component contributed substantially to the master list of evaluation questions and added product planning and product data collection phases to the TEE methodology. The requirements for consideration of team training, testing, and the ISD frame of reference further expanded the master list for both product and process evaluation.

Other procedural considerations involved efforts to insure the usability of the TEE system from the perspective of the TEE analyst and his interface with the documentation. To facilitate the TEE process, procedures were developed to narrow the scope of an investigation to a manageable effort. TEE training materials and job aids were developed to assist the TEE analyst and his data collectors. Additionally, a separate data collector's manual was developed for personnel who are unskilled in training evaluation technology, but who can be expected to assist the TEE analyst.

The resulting TEE procedures are described in detail in Section V. Appendix B provides short forms of the master list of evaluation questions.

The discussion of procedural considerations thus far has focused on their incorporation as features in the TEE methodology. Of equal importance, however, was the procedural approach to developing the TEE system . From the previous discussion, it is evident that the literature contributed greatly to TEE development. The search for baseline methodologies and contributing sources in the literature was facilitated by access to interactive, on-line information retrieval services. TEE methodology development also benefited from the expert judgment of three instructional psychologists and inputs from the Army users of the TEE methodology. Finally, the technique of formative evaluation was employed. A series of tryouts was used to evaluate the methodological components and the user documentation during various stages of development. The tryouts were conducted using both instructional psychologists and Army Air Defense personnel. Air Defense training programs were evaluated in both product and process modes with the purpose of providing inputs to TEE revisions. A detailed discussion of the tryouts is contained in Section IV.

Events in TEE Methodology Development

Development of the TEE methodology began with a review of the literature on product evaluation, process evaluation, and team training. The results of this review are described in detail in Section III of this report.

The concepts and procedures of product and process evaluation were further explored in relation to overall ARI objectives and ongoing ARI developments, such as the evaluation and revision methodology development occurring at ARI Field Unit Ft. Knox, and the conduct of XM-1 and other TEEs using derivatives of the Harless Guidelines.

Following the preparation of a working paper describing the technical approach to methodology development (described earlier in this section), and the completion of the literature review, an outline of TEE procedures was developed. The outline identified issues to be addressed in relation to product and process evaluation and suggested corresponding components for the TEE methodology. Major topics included procedures for planning the TEE, categories of data to be collected, procedures for summarizing the data, and procedures for identifying and rating performance deficiencies. Based on this outline, the design of specific TEE procedures and worksheets commenced.

In the meantime, sections of the original Harless Guidelines were tried out on the Vulcan training program at Ft. Bliss. Problems in using the worksheets or understanding the intent of specific items on the worksheets were encountered and noted for revision. Section IV contains a discussion of the tryout of the Harless Guidelines on Vulcan training. Greater detail on the formative evaluation and a description of components and procedures of the final iteration of the TEE system are contained in Sections IV and V of this report.

As a consequence of validating the original Harless Guidelines, it was determined that two "user's guides" would need to be designed:

a. A "TEE Evaluator's Handbook" for the TEE analyst, containing all TEE procedures and training materials related to both product and process evaluation.

b. A "Data Collector's Manual" containing a subset of references and training materials relevant to the data collection phase of process evaluation only.

A preliminary TEE Evaluator's Handbook was produced, including procedures for conducting the TEE, 18 worksheets with guidance for their use, a glossary of TEE terminology, a master list of 93 TEE questions, and data collector training materials.

Several sections of the user's guides relevant to product evaluation were then tried out using course materials from SGT York OT training packages. Results of the TEE tryout indicated that while the basic TEE components appeared to be sound, work needed to be done in transforming the procedures into an easily understood user-oriented format. Specific deficiencies were then identified and revisions were incorporated into the user's guides in the following areas: a. the master list of TEE questions; b. the associated rating scales; and c. the data collection formats. Product and process job performance aids (JPA) were also designed to provide ready reference for the TEE analyst and data collectors, respectively. These JPAs contain detailed guidance for making ratings on selected TEE questions. Finally, procedures for summarizing and interpreting the data were developed. More detail on the tryout of the TEE user's guides on SGT York OT training is provided in Section IV.

A formative evaluation of both product and process TEE procedures was then conducted on Nike Hercules training at Ft. Bliss. This tryout involved indoctrination and training of TEE analysts and data collectors, as well as a sampling of the TEE procedures applicable to a segment of an ongoing training schedule. The effort focussed initially on the product TEE and then followed through with a process TEE on the same curriculum. Comments from all TEE participants and an ARI debriefing contributed to subsequent TEE revisions aimed mainly at reducing complexities in the procedures whenever possible.

SECTION III

Literature Base For The TEE Methodology

Sources Searched

Early in the TEE methodology development process, it was recognized that three sources had the greatest potential for providing major inputs to the TEE methodology. The first one, Guidelines for Conducting a Training Program Evaluation (Harless Performance Guild, Note 1), was referenced by ARI as a foundation document. It was thoroughly reviewed, and many parts of it were ultimately included in the TEE methodology in some form. The second major source was The Instructional Quality Inventory (Ellis & Wulfeck, 1978; Ellis, Wulfeck, and Fredericks, 1979), a methodology useful for evaluating the adequacy of training materials. The IQI was also reviewed in detail and provided substantial inputs to the TEE methodology. The Interservice Procedures for Instructional Systems Development (Branson, Rayner, Cox, Furman, King, and Hannum, 1975), was a third source. The IPISD served as a useful general reference, providing guidance in many areas as the methodology developed.

Additional sources were sought through literature searches of the Educational Resources Information Center, the National Technical Information Service, and the Defense Technical Information Center. These searches yielded approximately 1050 abstracts in the areas of training, evaluation, and training effectiveness, as well as 490 abstracts in the areas of unit, team, or collective training. From these abstracts 52 documents were selected for review. A total of 20 documents eventually provided specific inputs to the TEE methodology or provided guidance on their development. These documents and their relationship to the methodology are reviewed in the following paragraphs. A more comprehensive review of the literature is contained in our selection of abstracts relevant to training evaluation and team training in Appendix A.

The Baseline for the TEE Methodology

Guidelines for Conducting a Training Program Evaluation. The Harless Performance Guild (Note 1) provides a method for observing the process of instructor-led training. It is a job aid system consisting of a set of guidelines on how to conduct an evaluation, plus 19 worksheets keyed to the Harless Guidelines.

Essentially, there is one worksheet for each of the tasks of the evaluation. Each worksheet is for the purpose of either collecting data, summarizing data, or interpreting data. The Harless Guidelines give step-by-step directions and examples for using the worksheets to conduct a training effectiveness analysis.

The approach taken for the evaluation is built around five "phases":

a. Phase A: Plan the TEA.

b. Phase B: Observe training and testing.

c. Phase C: Assess quality of trainee performance.

c. Phase D: Hypothesize and investigate training-causes of deficiencies.

d. Phase E: Document findings of the TEA.

The scope of the system extends from the initial request to perform an evaluation on some Army course to a report outlining the results of the investigation. The scope of an evaluation project includes collecting data for the purposes of identifying "failures" in the performance of tasks by trainees (called "performance deficiencies" in the Harless Guidelines). The scope includes methods for determining if these deficiencies were probably caused by a deficiency in training received or in training management and student selection problems.

Each Harless phase is divided into a number of tasks. Each phase and task is described below.

The purpose of Phase A, plan the TEA, is to collect as much background and logistical information as feasible in advance of observing the training and testing that will be analyzed. This involves assembly of documentation and plans that already exist, and an initial meeting with the requestor/user of the TEA.

No actual judgment of the training or materials is involved in Phase A. The effort primarily involves getting ready to perform the TEA. Phase A tasks are:

a. Request background information and an initial meeting with the training organization.

b. Become familiar with the technical content of the performance tasks to be trained and the training to be conducted. Describe the performance tasks relevant to the training.

c. Select performance tasks to be evaluated in the TEA project.

d. Describe the training events, purpose, and materials of the training.

e. Prepare a schedule of events for the TEA.

During Phase B, the training and testing events selected in Phase A are directly observed. The primary purpose of Phase B observations is to gather information that will be useful later to analyze the results of the training.

If the trainees are unable to meet performance criteria when they engage in operational tests, the information collected in Phase B will be useful in determining the causes of the performance deficiencies.

If the trainees are able to perform to criteria, the observation data collected in Phase B will serve as a record of the training and testing conditions that produced the result. A record of success is often as helpful to the designers and deliverers of training as feedback on failures. The tasks of Phase B are:

a. Prepare for observations.

b. Describe trainee characteristics.

c. Describe instructor characteristics.

d. Characterize training environment.

e. Observe input-training events.

f. Observe integrated practice and testing events.

g. Collect trainee reactions.

h. Collect instructor reactions.

Phase C, assess quality of trainee performance, has two major purposes:

a. Make assessments of how well the trainees can perform the tasks selected on completion of training.

b. If performance criteria not met, select tasks for further analysis of probable causes of the deficiencies.

The tasks of Phase C are:

a. Summarize raw test data.

b. Analyze test data.

c. Select performance deficiencies for further analysis.

The guidelines in Phase D of the TEA, hypothesize and investigate training-causes of deficiencies, are very much like a problem-solving process based on the scientific method. The performance deficiencies have been defined in previous phases of the TEA. Attention now turns to one of the possible causes of the deficiencies.

Implied in the above is the assumption that:

a. There may be causes of performance problems (deficiencies) other than the lack of skill/knowledge on the part of the trainee.

b. Any given performance deficiency may exist due to a combination of reasons (multiple causes).

In Phase D the attempt is made to sort the performance deficiencies caused by a lack of skills or knowledge from those that have other primary causes. The tasks in Phase D are:

a. Match evidence for and against hypothesis.

b. Interpret findings.

The purpose of Phase E, document TEA, is to summarize the findings of the TEA and to prepare a report documenting the effort. There are two tasks:

a. Review all data and analysis.

b. Outline and prepare report.

The procedures outlined for the above phases were generally incorporated in the development of the TEE methodology. In the planning phase, A, the Harless approach was found to be adequate and only a few modifications and additions were made based on our experience in a tryout of Harless' Guidelines on Air Defense training. For example, the need to discuss the purpose of the TEE in the initial meeting was noted; the documenting of task conditions and standards was deferred until after tasks are selected for evaluation (to save time), and some of the worksheets were modified or replaced with actual course documentation.

Phase B on data collection was also deemed usable for process evaluation. The first three tasks, on preparing for observations and collecting trainee and instructor characteristics, were included with

only minor modifications. The next three tasks, which deal with the observation of different types of training events, were not used as procedural steps in the TEE; however, they provided roughly one-third of the evaluation questions used in the TEE methodology. The data collection procedures have been largely revised, including training materials and procedures for training data collectors. The final two tasks on the collection of trainee and instructor reactions were found to be somewhat general in nature and were replaced by a procedure which is much more specific.

The data summary phase (C) procedures were found useful for summarizing performance test data. Those parts were included with only minor conceptual modifications. For example, only the first test trial will be summarized in a TEE, while Harless worksheets call for data from three trials. The TEE data summary and analysis procedures go beyond Harless' methods in providing a more organized set of guidelines and worksheets for summarizing the actual training and testing problems that were discovered in the TEE for each task. The TEE procedures then allow the analyst to make judgments on the seriousness of the problems for each task in terms of their probable impact on test adequacy and student performance. These judgments are then combined with task performance data in such a way that the most valid and reliable data are used in specifying task "discrepancies" (similar to Harless' deficiencies). Task 3 on selecting performance deficiencies for further analysis was omitted. It is felt that this task is more appropriate as the first step of a separate revision methodology to be developed in a follow-on effort to this contract, since it is applicable only when revisions are to be made, not when the purpose of the evaluation is merely to evaluate the course.

Harless' Phase D examines each performance deficiency and attempts to judge whether the students actually are deficient in terms of their skills or knowledge. Otherwise the deficiency is attributed to problems that would cause low test scores even though the students could actually perform the tasks. In the TEE methodology however, data is collected on all of the potential problems, including a product and process evaluation of the tests. Furthermore, test adequacy is judged separately from training adequacy in the TEE data analysis process and then combined with other data, as described above. That process replaces the one in Harless' Phase D and results in a final rating that already includes the consideration of whether or not a deficiency is a skill/knowledge deficiency. The only factors that are not considered relate to whether training devices or real equipment used in training or testing are designed such that they reduce proficiency in task performance. Equipment related issues have also been deferred to a separate methodology development effort.

Phase E, document TEA, was deemed appropriate, but guidance on interpreting and displaying data has been added along with suggestions on preserving the TEE documentation for later use.

The Instructional Quality Inventory. The IQI (Ellis and Wulfeck, 1978) is a methodology for evaluating existing training materials. It served as the baseline for the TEE product evaluation methods and presents a basic evaluation philosophy which was generally adhered to in developing the TEE methodology. This philosophy constitutes an approach which has six major components:

a. Purpose-objective consistency. The first set of IQI diagnoses is concerned with determining whether or not each objective is one which should be taught. This justification of the objectives requires four steps. First, analyze the purpose of the lesson to be taught and classify it on the basis of important characteristics. Second, analyze the objectives and classify them on the basis of the same characteristics. Third, compare the classification of each objective with the classification of the purpose. If they are not the same, the objective should be revised to be consistent with the purpose. Finally, make sure that no important objectives have been left out.

b. Objective adequacy. Once it has been determined that the objectives for a lesson are consistent with the purpose of the lesson, the second set of IQI diagnoses is intended to determine whether or not each objective is adequate. Three important criteria of objective adequacy must be clearly specified: a. the desired student behavior; b. the conditions under which the behavior is to be performed; and c. the standards for the acceptable performance of the behavior.

c. Objective - test consistency. Having determined that the objectives for a lesson are justified and adequate, the third set of IQI diagnoses is intended to determine whether or not the test items are consistent with those objectives. This analysis requires four steps. First, classify the objectives on the basis of important characteristics. Second, classify each test time on the basis of those same characteristics. Third, match each test item with the objective it tests (if any), compare their classifications, and (if necessary) revise the test item to be consistent with the objective. Finally, make sure all the objectives are tested.

d. Test adequacy. Once it has been determined that the test items for a lesson are consistent with the justified objectives, the fourth set of diagnoses is intended to determine whether or not the test items are adequate. There are two important aspects of test adequacy that have received considerable attention: a. the reliability of test items, and b. the technical correctness of the format of each test item. The IQI calls for the analysis of these aspects of test adequacy.

There are some other aspects of test adequacy that have been largely overlooked or have received considerably less attention. These aspects are of two types: a. those that apply to the adequacy of single test items, such as some characteristics of the information provided and of the behavior required, and b. those that relate to sets of test items, such as item sampling, item sequencing, and criterion-level determination. These aspects of test adequacy are also considered in the IQI.

e. Test - presentation consistency. Having determined that the test for a lesson is consistent with the justified objectives and is adequate with respect to reliability, item format, and other aspects of quality, the fifth set of diagnoses is intended to determine whether or not an instructional presentation is consistent with its test item(s) - that is, to determine whether or not the presentation contains the information necessary for the student to learn how to perform as required by the test.

This analysis requires three steps. First, determine the task level of the test item(s) on an objective. Second, determine what the presentation needs to contain in order to be able to teach at that task level. And finally, analyze the presentation to see whether or not it contains those components and only those components. If it does, the presentation is consistent with its corresponding test item(s). If not, the presentation should be revised.

f. Presentation adequacy. Having made sure that each instructional presentation contains the appropriate primary presentation forms for teaching at the desired task level (test - presentation consistency), one can go to the sixth and last set of IQI diagnoses, which is intended to determine whether or not each primary presentation form is accompanied by the necessary secondary presentation forms and has the necessary strategy components and characteristics to teach well at the desired task level (presentation adequacy). There are two major aspects of presentation adequacy: a. what strategy components should be included in and with each primary presentation form, and b. what characteristics each of those strategy components should have.

The above approach is illustrated in Figure 1. In addition to providing this basic approach, the IQI supplied inputs to nearly half of the evaluation questions used in the TEE methodology, as well as some statements, examples, and practice items from training materials on content type and task level (Ellis, et al., 1979).

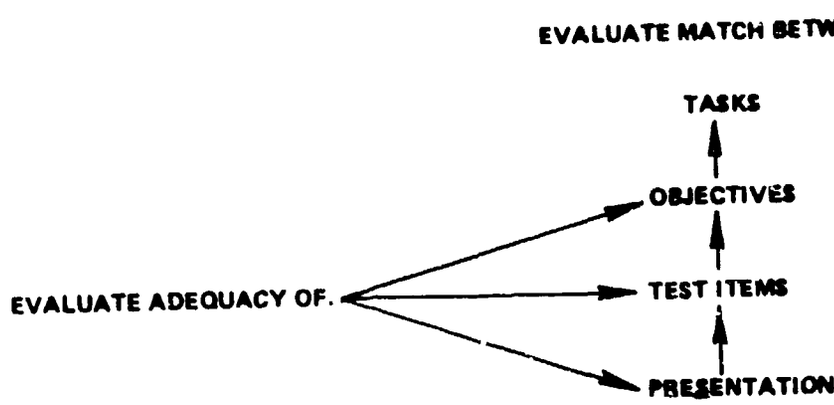

Figure 1. General evaluation strategy followed in the IQI and the TEE Methodology.

The Interservice Procedures for Instructional Systems Development. The IPISD (Branson, et al., 1975) are a complete set of guidelines on the development, production, and evaluation of instructional materials. Besides following the basic philosophy of the IQI, the TEE methodology also follows the structure of IPISD block V.I: conduct internal evaluation (Figure 2). Each of the general steps outlined there is or can be conducted using the much more detailed specifications in the TEE methodology. There is one major difference involving the nature of the ISD "progress" evaluation plan. Using that plan, the adequacy of training materials is evaluated during the ISD development effort by checking with other training development personnel to ensure that each of the ISD steps has been conducted. In a TEE, there is no guarantee that the complete ISD model will have been used, if at all. Even when it has been used, it may be difficult to locate all of the development personnel, depending on how long it has been since the course was developed. Therefore, the TEE product evaluation is conducted instead of the ISD progress evaluation. In the TEE, the training materials themselves are examined, along with their corresponding tasks or objectives to insure that they adhere to ISD principles.

Several basic concepts were also drawn from IPISD, e. g., tasks, terminal learning objectives (TLO), learning objectives (LO), and entry skills. The ISD model provided specific inputs to evaluation questions on entry skills and the sequencing of objectives.

Other Contributions from the Literature

Besides the three major sources described above, fifteen other references provided ideas and inputs used in the TEE methodology. Each of these and its relationship to the methodology is described below.

A list of evaluation questions dealing with objectives, tests, and the training process was obtained from the ARI Field Unit at Fort Knox (Kristiansen, Note 2). These questions were found to be a valid evaluation and to have a great deal of overlap with the material in the IQI. They did, however, provide the basis for a number of additional TEE evaluation questions.

Schulz and Farrel (1980) developed a set of job aid manuals and accompanying job performance aids which contain detailed specifications on how to develop instruction following the Interservice Procedures for Instructional Systems Development (Branson, et al., 1975). Segments of

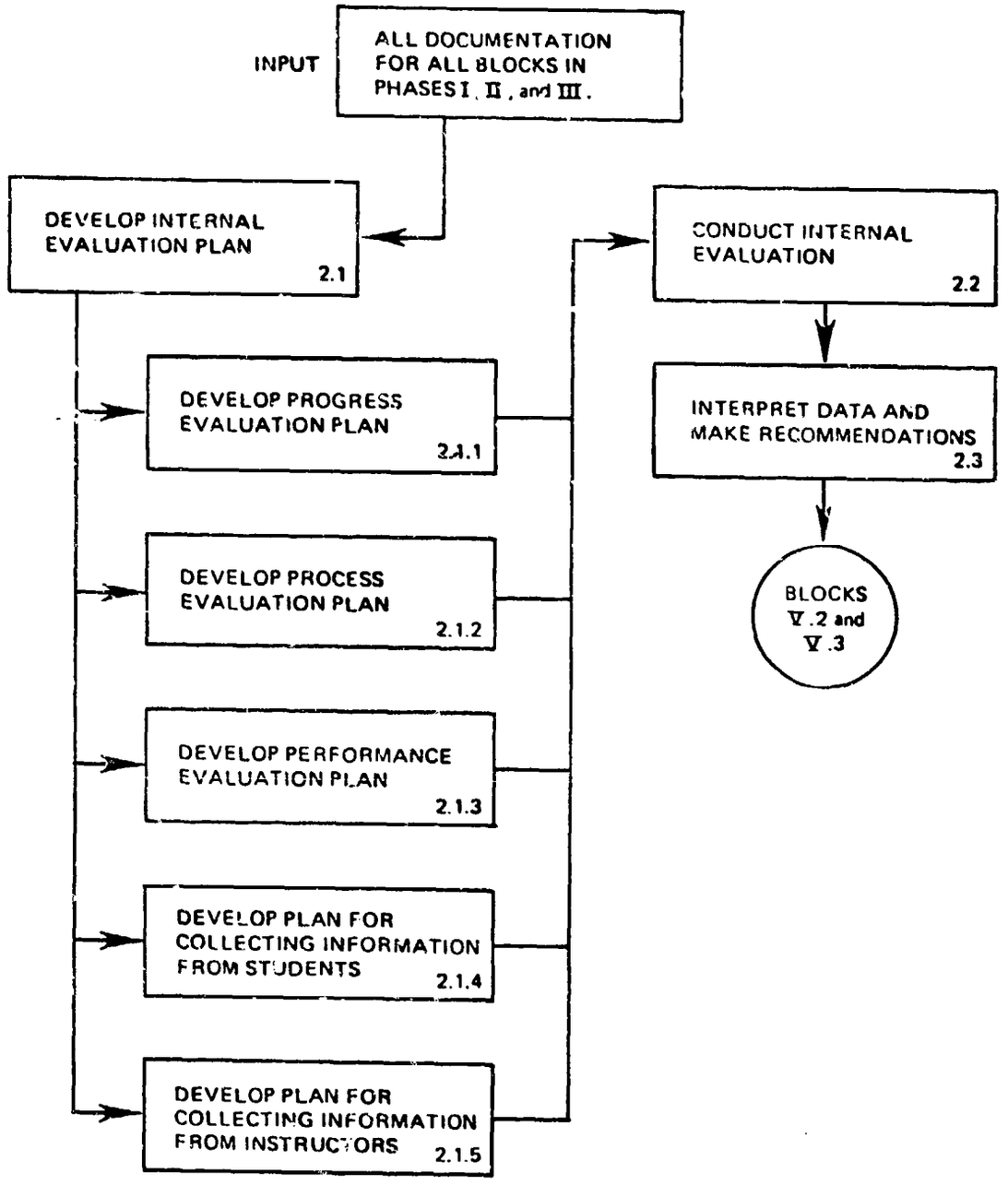

Figure 2. Flowchart of IPISD block V.1, conduct internal evaluation. (From Branson,et al.,1975,Phases IV & V, p. 62.)

these on indicator behaviors, critical task elements, and the clarity of narration contributed to a few TEE evaluation questions.

A large portion of the TEE evaluation questions dealing with tests came from four sources:

a. Guidebook for Developing Criterion-referenced Tests (Swezey and Pearlstein, 1975).

b. A Handbook on Item Writing for Criterion-referenced Tests (Roid and Haladyna, Note 3).

c. Writing Technically Correct Test Items (Courseware, Note 4).

d. Tests, Measurement of Student Achievement (U.S. Navy, 1976).

The first three sources are useful sets of methods and criteria on developing tests and test items. The last source is a U. S. Navy data item description giving requirements for test items in several different formats. These sources have some similarities among themselves and with the IQI; however, they each contributed several unique test and test item criteria to the TEE methodology.

The IQI incorporates a basic instructional strategy framework in its evaluation questions. Its components include statements, examples, practice, and feedback. Carey and Briggs (1977) provide a more comprehensive framework with which instructional strategies can be compared using Gagne's events of instruction (pp. 275-277). Table 1 shows the relationship between the systems. Evaluation questions were included in the TEE methodology which correspond to the first three events shown in the table, because they have no corresponding counterparts in the IQI. Enhancing retention and transfer is the only other event not covered. Carey and Briggs give strategies for this event that primarily involve additional practice after the performance has been assessed. Since such assessments (tests) are usually given only once during a course and the scope of a TEE extends no further than the final test, this event was deemed not applicable.

Table 1

Relationship Between Gagne's Events of Instruction and the IQI's Presentation Components

GAGNE	IQI
Gaining Attention	
Informing Learner of the Objective	
Stimulating Recall of Prerequisites	
Presenting the Stimulus Material	Statements and Examples
Providing Learning Guidance	Statements and Example Help
Eliciting the Performance	Practice
Providing Feedback	Feedback and Feedback Help
Assessing Performance	(Testing)
Enhancing Retention and and Transfer	

Dick and Carey (1978) present a complete instructional design model. A small portion of this, on preinstructional activity (pp. 106 107), describes certain attention-getting and motivational techniques included in one TEE evaluation question.

Dick (Note 5) describes a method for evaluating instruction which is being developed. It is done by expert review or some form of empirical validation appropriate to each stage of the design process. As one part of this, he explains a method for empirically verifying a learning hierarchy. It involves recording the percentage of students passing each of the objectives for a task/TLO on a learning hierarchy and comparing the performance relationships among superordinate and subordinate objectives. A very similar technique is used to analyze such relationships in the TEE methodology.

TEE evaluation questions dealing with the technical quality of written material and how easy it is to understand were taken mostly from the Guidebook for the Development of Army Training Literature (Kern, Sticht, Welty, and Hauke, 1976). This excellent guide was designed as a resource for authors of performance-oriented training.

An article by Champagne and Klopfer (1974) contributed to evaluation questions on the technical quality of audio-visual materials and the completeness and appropriateness of course administration directions. The article explains a formative evaluation methodology for science curriculum and gives several lists of evaluation questions oriented to different aspects of the curriculum, i.e., planning, student materials, student behavior, instructor's materials, and the marketability of the curriculum.

Shriver (1975) describes a thorough method for developing fully proceduralized job performance aids. It provided adequacy criteria to an evaluation question on the adequacy of job performance aids used in training.

Braby, Kincaid, and Aagard (1978) contribute examples of types of memory aids for one evaluation question. This interesting report gives guidelines for deciding when to use mnemonics, what types to use, and how to use nine mnemonic techniques.

Two sources provided some basic guidance on the evaluation of team training. Thurmond (1980) presents a study exploring the extension of the IPISD model to team training and its integration into the Army training acquisition and implementation system, the life cycle systems management model. Wagner, Nibbits, Rosenblat, and Schulz (1977) provides a comprehensive literature review describing existing instructional and evaluative techniques applicable to team training. Gaps in current knowledge are identified for future research.

Three team training variables are embodied in the TEE evaluation questions dealing with team training: practice, simulation fidelity, and feedback. These are cited as among the most important factors influencing team training outcomes. (Thurmond, 1980: Wagner, et al., 1977).

SECTION IV

Formative Evaluation Of The TEE Methodology

Section II outlines the conceptual and procedural considerations and the events involved in developing the TEE methodology. Section III discusses the literature base for the TEE methodology. The purpose of Section IV is to describe the formative evaluation process and the corresponding revisions that ultimately led to the TEE system presented in Section V.

The formative evaluation process, which contributed to the development of the TEE methodology and supporting documentation, consisted of three primary events: validation of the existing Harless Guidelines; tryout of the preliminary TEE methodology for product evalution; and tryout of the revised TEE methodology for both product and process evaluation. These three events are described in the following two subsections. Prior to reviewing these events, however, it is worthwhile to note that the initial plan was to conduct the tryouts on training for operational tests. Training packages for the SGT York OT were to provide the basis for validating the Harless Guidelines, while PATRIOT training packages were to provide the basis for tryout of the revised TEE methodology. However, it was not possible to apply the Harless Guidelines to the SGT York training packages apart from their implementation, since the methodology was process evaluation oriented and lacking a product evaluation component. Administrative reasons further precluded the observation of either SGT York or PATRIOT training for a test of the Harless Guidelines or the revised TEE methodology. This problem was ultimately solved by the substitution of on-going training on the Vulcan and Nike Hercules systems as the basis for methodological tryouts.

Validation of the Harless Guidelines

During the period of 13-22 January 1981, applicable portions of the Harless Guidelines were used to evaluate a segment of the on-going Vulcan training curriculum at Fort Bliss. Four evaluators participated in the exercise: one instructional psychologist from the contractor's team; one psychologist from ARI; and two Army officers who served as subject matter experts as well as data collectors and training analysts. These evaluators observed both classroom and hands-on training as conducted in accordance with the Vulcan soldier's manual and the Vulcan POI.

The exercise began with the planning procedures as prescribed by the Harless Guidelines. This phase of the methodology also provided the forum for indoctrinating the evaluators in the purpose and procedures of the Harless Guidelines. As expected, planning required tailoring of specific procedures, but the overall framework was found to be usable. The most demanding step in the planning procedures was that of task selection. Further definition of classification factors appeared to be needed to clarify the process. In general, much of the prescribed planning documentation was found to be unnecessary in this limited exercise, since the requirement for flexibility became the rule rather than the exception.

Data collection was based on observations of each category of training event as defined by the Harless Guidelines, with the exception of integrated

practice and integrated testing. Revisions were made to the corresponding worksheets based on the judgments of the evaluators. Four revised worksheets emerged from their efforts: a general training observations form; a classroom instruction observations form; a demonstration or practice observations form; and a general task rating form. Additionally, a trainee opinion questionnaire was developed. These worksheets were subsequently validated through data collection activities over the remainder of the Vulcan tryout.

Tryout of the TEE Methodology

Product TEE. Two training packages developed by contractors for the SGT York OT provided the basis for a tryout of the preliminary TEE methodology in the product evaluation mode. Both training packages reflected a systems approach to development, but the two differed markedly in the nature of the documentation. Differences in the presentation of task statements and learning objectives had the greatest effect on the applicability of the TEE methodology, with only one of the two training packages appearing adequate for purposes of exercising the product evaluation methodology. Two instructional psychologists served as TEE analysts for the tryout.

The first SGT York training package to be used as a basis for a tryout of the TEE methodology contained the following documentation:

a. A front-end analysis.
b. Operator course student information sheets.
c. Lesson guides.
d. Audio-visual aids.
e. An operator's manual.
f. A student training record.
g. Tests.

While this training package appeared to be the more complete of the two for TEE tryout purposes, the opposite conclusion was ultimately reached. The problem was that the front-end analysis documentation had been prepared to support operator and maintainer training after system implementation, rather than training for the OT. Therefore, the relationship of the tasks and training objectives to OT training could not be determined. While the TEE methodology could have been applied further to this training package, the effort would have required extensive interaction with subject matter experts.

The second SGT York training package, to which the TEE methodology was applied for a product evaluation tryout, consisted primarily of a single manual with the following six sections:

a. Training concept.
b. Course outline.
c. Lesson guides.
d. Training devices.
e. Government furnished equipment list.
f. References.

The course outline and the detailed lesson guides provided the most useful documentation for TEE tryout purposes. Tasks were specified in the course materials, but no conditions or standards were documented. Nevertheless, a sample of complete task statements was derived. These tasks became the terminal learning objectives which were traced throughout the training package using the TEE methodology. The tryout proceeded through the planning phase of the TEE methodology with few problems. However, the need for the system to allow a high degree of flexibility in tailoring the TEE procedures to the unique requirements of each training package was reinforced. Phase B, data collection for product evaluation, followed. It was here that specific revision requirements emerged in three areas. First, the master list of evaluation questions was found to be in need of semantic revisions for greater clarity to the user. Second, the evaluation scales were found to be too complex and in need of conceptual refinement. Third, formats for data collection worksheets were found to be less than optimum in facilitating the users' efforts.

Major revisions were made to the TEE methodology as a result of: a. the product TEE tryouts; b. the previous validation of the process oriented Harless Guidelines; and c. the judgments of a team of instructional psychologists. Revisions focusing primarily on the master list of evaluation questions, the rating scales, and the data collection formats, and secondarily on streamlining the TEE procedures were incorporated. Additionally, procedures for summarizing and interpreting the data were developed. The revised TEE methodology were then applied to Nike Hercules training for a more complete formative evaluation as described in the following subsection.

Product and Process TEE. Formative evaluation of the revised TEE methodology was conducted on Nike Hercules training at Fort Bliss during the period of June 8 through June 18, 1981. Psychologists representing the contractor and ARI served as TEE analysts, while two Army enlisted men (E-7s) with the Nike Hercules MOS served as data collectors. Phases A through D of the TEE methodology were exercised, thus providing coverage of TEE planning, product data collection, and process data collection on a single training program. An effort was made to conduct the exercise in a manner as close as possible to conditions expected to be present under actual TEE situations. Accordingly, attention to training the TEE participants was a major consideration.

The Nike Hercules training to which the TEE methodology was applied involved fifteen students in two classes. Two instructors were assigned to each class. The goal of the instruction was to familiarize the students with all tasks of the MOS at the battery level and to prepare them for more specific training in the field. Students were required to perform each function in the MOS to proficiency during training, but were not expected to retain proficiency at their assigned units without further practice. The course consisted of both classroom instruction and hands-on training using operational equipment.

The general approach to the TEE exercise can be characterized as follows:

a. Review of course documentation and evaluation/organization of objectives.

b. Description of instructional methods and specification of training events by type.

c. Evaluation of lesson plans, tests, training manuals, and operators' manuals.

d. Planning for process evaluation.

e. Outlining the training schedule and tailoring TEE worksheets for each training event.

f. Training the data collectors and conducting observations.

g. Administration of instructor and student attitude surveys.

h. Tabulation of the process data and analysis of both product and process data.

i. Formulation of conclusions and recommendations.

The ultimate objectives of the TEE were to identify discrepant tasks and to determine corresponding training deficiencies.

A number of specific TEE deficiencies were identified from the formative evaluation. These can be characterized generally as semantic or format problems in need of fine tuning. More general TEE deficiencies pointed to the need for further reductions in the complexity of procedures and the reading level required of the users. Additionally, the need for more support to the TEE analysts and data collectors suggested requirements for greater emphasis on TEE training and the incorporation of job performance aids.

Efforts to correct the TEE deficiencies identified during the formative evaluation were directed toward the TEE Evaluator's Handbook and the Data Collector's Manual prior to project completion. The resulting TEE methodology is discussed in Section V of this report.

SECTION V

The TEE System

The TEE system consists of the TEE Evaluator's Handbook, Guidelines for Conducting a Training Effectiveness Evaluation, a Data Collector's Manual, job performance aids for product and process evaluation, and a set of reproducible masters for worksheets and training materials.

The TEE Evaluator's Handbook includes the following major components:

a. Guidelines for conducting the TEE (Phases A through F).
b. A master list of evaluation questions.
c. Job aids for conducting product and process evaluations.
d. Training materials for TEE analysts and associate analysts.
e. A set of worksheets keyed to the guidelines.

Worksheets are used with 11 out of the 18 TEE tasks in conducting an evaluation. In other tasks, actual course documentation is annotated and used as a worksheet. Each worksheet is for the purpose of either planning to collect data, collecting data, summarizing data, or interpreting data. The Guidelines give step-by-step directions and examples for using the worksheets to conduct an evaluation.

The primary components of the Data Collector's Manual are:

a. The master list of evaluation questions edited for use by data collectors.

b. Training materials on how to classify task level and to recognize glossary terms.

The job performance aid for use in product evaluation is a subset of the master list of evaluation questions. It contains only those TEE questions appropriate to product evaluation along with abbreviated rating guidance.

The job performance aid for use in process evaluation is also a subset of the master list. It contains detailed guidance for those process TEE questions which cannot be rated without directions beyond the questions themselves.

The scope of the TEE methodology extends from the initial request to perform an evaluation on a designated Army course to a report outlining the results of the investigation. The scope of a TEE project includes collecting data for the purposes of identifying failures in performance of tasks by the trainees and failures in the instructional system itself. Data on both the training process, collected by actually observing the training, and on the training materials can be included.

Along these same lines, it may not always be necessary to conduct a "full TEE" using every task and step outlined in Figures 3 and 4. Some of

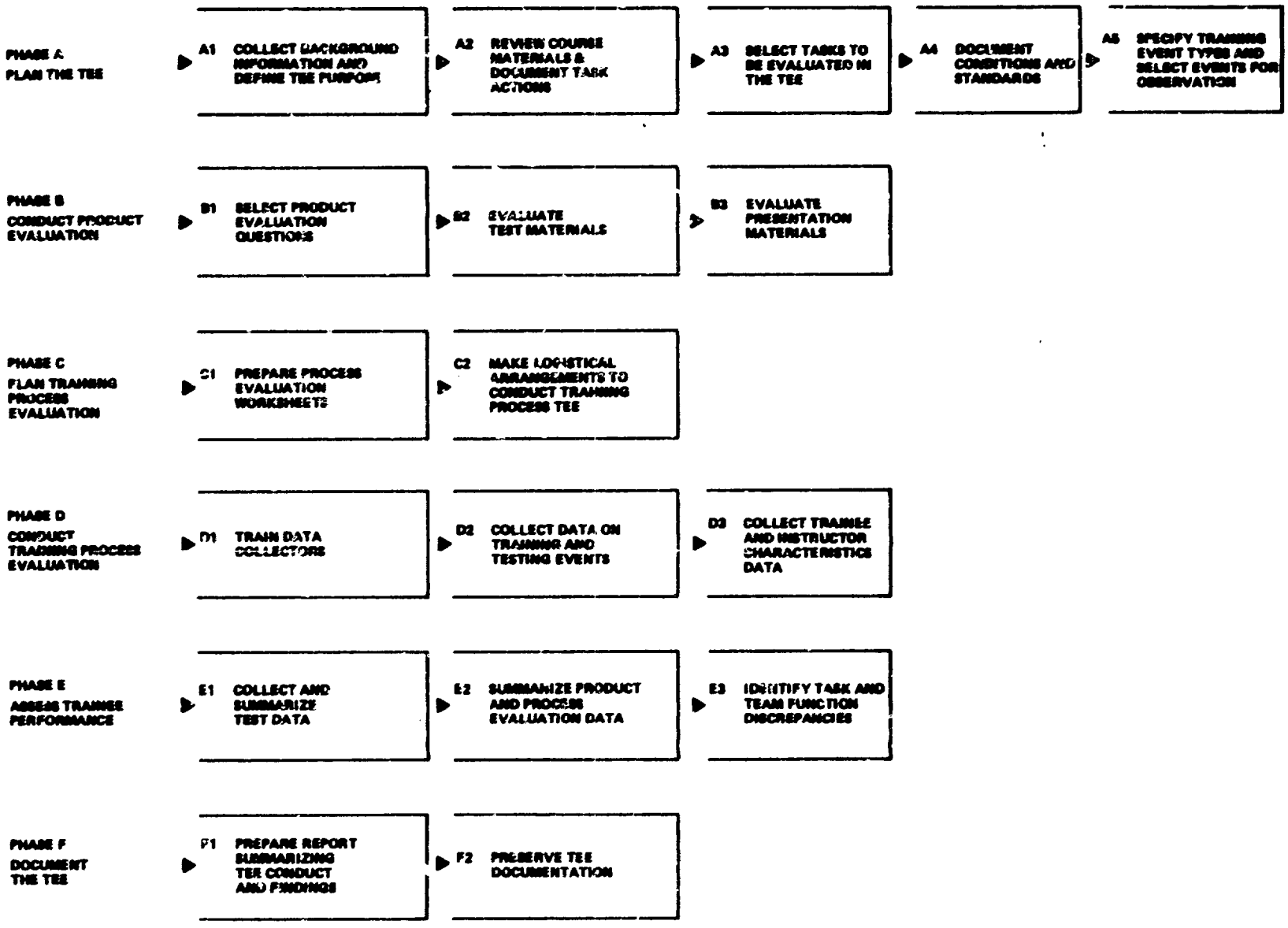

Figure 3. Conceptual model of TEE procedures.

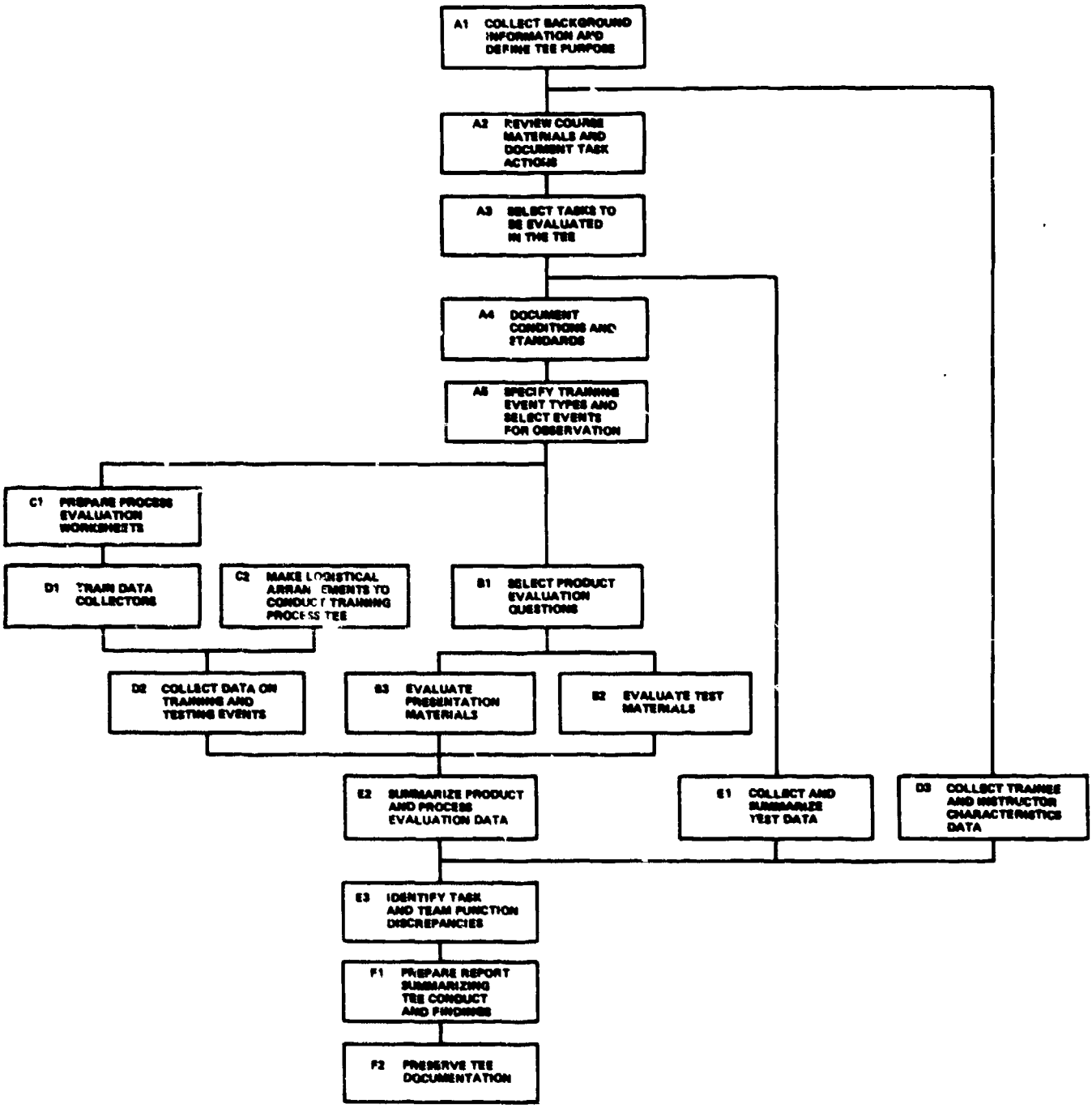

Figure 4. Procedural hierarchy of TEE tasks.

the data collection steps may be omitted when the information is already known, or one may wish to concentrate on certain aspects of the training which are known to be bad. A complete TEE may not be within the resources available.

A number of different types of personnel will be involved in the TEE. In addition to the sponsor, the organization that decides to conduct the TEE and utilizes the results, and the training organization which will be evaluated, there are three possible types of personnel who may also be involved:

a. The TEE analyst will be responsible for planning and conducting the TEE, analyzing the data, and writing the final report on the TEE. The analyst may need to supervise a team of data collectors, and should have some type of training development or evaluation background and should preferably have some familiarity with instructional systems development (e.g., see Branson, et al., 1975). Completion of an ISD workshop and/or a criterion referenced instruction course (or their equivalent) is recommended.

b. An associate analyst may be employed to assist the TEE analyst with some or all of his responsibilities, and should be a subject matter expert (SME) in the training to be evaluated. If a team of data collectors (described below) is employed, one of the associate analyst's primary responsibilities should be to supervise much of the data collection activity, acting as an interface between the TEE analyst and the data collectors. In that event, the associate analyst should be an Army officer, noncommissioned officer, or enlisted person with a higher rank than that of the data collectors selected for the TEE. The associate analyst need not have a training background, although that would be desirable.

c. Data collectors should be employed whenever the TEE analyst and an associate analyst cannot adequately observe the training alone. This may occur when there are several classes being taught at the same time. Data collectors can be any available personnel, although it is desirable for as many of them as possible to be subject matter experts in the training they will observe.

Another characteristic important in associate analysts and data collectors is reading ability. Without this quality they will be difficult to train and will not function well in the TEE. The associate analyst should have a reading grade level of at least 10. It is also preferable for data collectors to have a tenth grade reading ability. However, subject matter experts with a lower reading level may be paired with non-SMEs who have a tenth grade reading ability. In no case should data collectors be recruited with reading grade levels lower than 8.

Especially in the early stages of the TEE, the TEE analyst should become well acquainted with the sponsor (the organization requesting the TEE) in order to understand what is expected from the TEE and what resources will be available. It is also important that the channels of communication be opened, putting the analyst in direct contact with the training organization conducting the training to be evaluated.

The TEE methodology consists of six phases:

a. Phase A: Plan the TEE.
b. Phase B: Conduct product evaluation
c. Phase C: Plan training process evaluation.
d. Phase D: Conduct training process evaluation.
e. Phase E: Assess trainee performance
f. Phase F: Document the TEE.

Each of these phases is described below.

Phase A: Plan the TEE

The purpose of Phase A is to collect as much background and logistical information as feasible in advance of observing the training and testing that will be analyzed. This involves assembly of documentation and plans that already exist, and an initial meeting with the requestor/user of the TEE. It also involves decisions regarding the TEE's purpose and the selection of appropriate questions to ask during the TEE.

Little actual judgment of the training or materials is involved in Phase A. The effort primarily involves getting ready to perform the TEE. However it does involve a review of the tasks and objectives addressed in the course.

These are the tasks for Phase A:

Task A1: Collect background information and define TEE purpose. In this task, the TEE analyst requests a meeting with the organizations that are most knowledgeable of the training content and methods of the course that the TEE will analyze, and also requests the following materials and documentation relevant to the project:

a. Task documentation and/or job data worksheets.

b. ISD or other documentation relevant to the training events that will occur, such as objectives, hierarchies, lesson plans, and practical exercises.

c. Test administration directions and/or any documentation concerning the methods that will be used during and at the end of the course to evaluate the performance of the trainees.

d. Commander's manual relevant to training.

e. Soldier's manual(s) relevant to the training.

f. Course materials to be used by the trainees.

g. Testing/evaluation instruments to be used in the course.

h. Training/testing schedules.

i. Course administration policies.

This documentation is studied prior to the initial meeting with representatives of the training organization. A worksheet entitled "Background on the Project" is provided to organize relevant TEE background information and for use as an

agenda for the initial meeting. Guidelines for completing this worksheet are included in the TEE Evaluator's Handbook.

Another subtask in Task A1 involves consideration of the overall purpose of the TEE. A TEE may be conducted solely to evaluate the adequacy of the training; it can be done with with an eye toward revising the training; or both. The purpose of the TEE may affect the selection of TEE system components to be utilized.

Defining "purpose" another way, a TEE will almost always have one general purpose, which is to identify training deficiencies and discrepancies. Training deficiencies are problems with the instructional materials or methods which would be expected to have an adverse affect on student performance. Discrepancies are actual substandard student performances on final course tests for particular tasks trained.

Data on task deficiencies and discrepancies can be used in several ways which correspond to the purposes of evaluation and revision:

a. To certify training for an operational test (evaluate).

b. To determine the quality of training in an ongoing course (evaluate).

c. To determine areas in the course which need to be revised in order to improve instructional effectiveness or efficiency (revise).

d. A combination of these uses (both evaluate and revise).

At the initial meeting, the TEE analyst also considers the level of effort required to conduct a product evaluation, a process evaluation, or both.

Task A2: Review course materials and document task actions. The analyst next reviews all available documentation and the information accumulated via the background worksheet. If practical, he discusses the key training tasks to be taught with the instructors or other subject matter experts.

The analyst then develops an accurate list of tasks or objectives and/or team functions. A glossary, self-instruction and practice on classifying content types and task levels, and "Guidelines for Task/TLO Actions" are provided in the manual as supporting reference material for this effort.

If an existing list of tasks or objectives is unavailable or is of poor quality, the Guidelines provide for options of extensively revising the existing task list or producing a new one.

Task A3: Select tasks to be evaluated in the TEE. In this task, the analyst decides whether or not all of the tasks in the course can be evaluated. If doing so is beyond existing resources or is impractical,

tasks must be selected for evaluation. The list of tasks under consideration is iteratively narrowed down, based on considerations of whether each task is: already known by the majority of trainees; known to have had performance problems in the past; difficult to learn; performed frequently on the job; and other factors.

Task A4: Document conditions and standards. The purpose of this task is to document objectives so that they validly reference tasks as they are performed on the job. Up to this point, the analyst has examined tasks in terms of the actions the trainee will be expected to exhibit after training. Now conditions and standards for tasks/terminal learning objectives must be evaluated or added if not already documented.

"General Guidance on Objectives" and "Guidelines for Evaluating Objectives" are included in the TEE manual as an aid to identifying/writing complete TLOs which correspond to tasks, and learning objectives which match the major subtasks. Criteria are given for determining whether each objective (TLO or LO) is correctly stated, classifiable by task level and content type, and appropriate. Guidance is also provided for determining whether all of the required LOs for a TLO are present.

Task A5: Specify training event types and select events for observation. Using the documentation collected in previous Phase A tasks, the analyst now makes a list of the types of training events that are employed in the course. These may include:

a. Classroom instruction.
b. Demonstration.
c. Practice.
d. Performance test.
e. Written test.
f. Oral test.
g. Integrated practice or test.
h. Individual study.
i. Help session.

The analyst then lists the training events associated with each lesson topic, annotates this list with the corresponding task numbers, and eliminates events for tasks not selected for evaluation.

Phase B: Conduct Product Evaluation

The purpose of Phase B is to evaluate the course materials and note deficiencies that are likely to cause performance problems on the final test. The general evaluation strategy is to check to see that test items are adequate and match the objectives, and to check the planned presentation for adequacy and its match to the test items.

These are the tasks for Phase B:

Task B1: Select product evaluation questions. In this task the master list

of evaluation questions has its first use. A list of the "short forms" of all master list questions, along with their rating scales, is shown in Appendix B. The master list is divided into two sections: one for tests, and one for the presentation of the instruction. In Task B1, the analyst identifies those questions that are applicable to a product evaluation (conducted using the course materials), and to the training situation. A similar process is conducted for process evaluation in Task C1.

Task B2: Evaluate test materials. In this task, the TEE analyst, referring to both the list of objectives and the course test(s), identifies the test questions that test each objective, then classifies each test question by task level and content type, eliminating inapplicable test items. Finally the TEE analyst conducts the evaluation of applicable test materials by asking each of the test-related master list questions (not previously eliminated) of each test or test item, as appropriate. A job performance aid for product evaluation is provided to facilitate this process.

Task B3: Evaluate presentation materials. The first step in this task is to gather the necessary course documentation:

a. The course objectives.
b. All lesson materials and manuals used by the students.
c. Any audio-visual equipment needed to hear or view the materials.
d. Lesson and course administrative directions.

The product evaluation JPA and the course outline with events to be evaluated are also required for this task. The analyst conducts the evaluation of presentation materials by asking each question in the presentation section of the master list, of each lesson, objective, presentation component (e.g., examples or practice), or the course as a whole, depending upon the level to which each question applies.

Phase C: Plan Training Process Evaluation.

The purpose of Phase C is to prepare a set of worksheets for recording training process observations appropriate to the TEE setting and to make plans to observe specific training events. General worksheets containing all possible TEE questions for different event types are included in an appendix to the TEE Evaluator's Handbook.

No actual judgment of the training or materials is involved in Phase C. The effort primarily involves getting ready to train data collectors and preparations for observing the training as it is being conducted in Phase D.

These are the tasks for Phase C:

Task C1: Prepare process evaluation worksheets. In this task the analyst locates the appropriate TEE worksheets for the training events

to be evaluated in the process mode. The general heading of each worksheet is modified to fit the analyst's particular requirements. TEE questions which are inapplicable to the events under consideration are eliminated from the worksheets. The analyst also has the option of creating his own process evaluation worksheets, tailored to the training settings to be observed.

The steps remaining in this task concern the preparation of trainee and instructor reaction instruments. While trainees are not particularly qualified to make subjective judgments about the quality of training, they are quite capable of observing what happened in the training and stating how it affected either the way they learned the material or their motivation for learning it. Therefore, trainee reactions which pinpoint individual learning problems or points in the instruction that substantially reduced motivation for learning are appropriate for TEE data collection.

Task C2: Make logistical arrangements to conduct the training process TEE. At this point, the TEE analyst must consider how many training events to observe. As a minimum, the following events must be observed for each task selected for evaluation:

a. A final test for each task selected for evaluation.

b. At least 50% of the demonstrations for each instructor involved.

c. As much practice as is practical for each task selected.

d. As much classroom instruction as possible.

e. As many of the other training event types as possible.

When the purpose of the TEE is to certify training prior to an operational test of a developing weapons system, it is more important to observe a large percentage or all of the training events for their entire duration. On the other hand, when ongoing training is observed that appears to be functioning fairly well and which is consistent across instructors, complete observation of every training event becomes less important.

After deciding which events to observe, the analyst prepares an "observation plan" or TEE schedule, and communicates his plans for an on-site visit to the personnel in charge of the

training to be observed. At this time the analyst obtains and reviews all training material and tests for the events to be evaluated, if this has not already been done.

Phase D: Conduct Training Process Evaluation.

During Phase D, the training and testing events selected in Phase A are observed directly. The primary purpose of Phase D observations is to gather information that will be useful in analyzing the results of the training, i.e., the performance scores.

If the trainees are unable to meet performance criteria when they engage in operational tests, the information collected in Phase D will be useful in determining the causes of the performance deficiencies.

If the trainees are able to perform to criteria, the observation data collected in Phase D will serve as a record of the training and testing conditions that produced the result. A record of success is often as helpful to the designers and deliverers of training as feedback on failures.

Three major types of information are collected during this phase:

a. Direct observation data.

b. Trainee and instructor characteristics.

c. Trainee and instructor reaction data.

These are the tasks for Phase D:

Task D1: Train data collectors. In this task the TEE analyst familiarizes data collectors with the worksheets they will use in observing the training and trains them in the skills necessary to do so.

The analyst first insures that there is a sufficient number of data collectors, if indeed the course has sufficient enrollment to require additional personnel to collect data. The analyst attempts to employ personnel who are subject matter experts in the training area to be observed and insures that at least some of them have a reading level of grade 10 or higher.

In the next few steps the analyst gathers the appropriate training materials, gives the data collectors an overview of the TEE process and the mechanics of what they will be doing, conducts training on terminology used in the observation worksheets that may be unfamiliar to the data collectors, and administers a self-instructional module on how to classify task level, a skill the data collectors will need in answering some of the questions on the worksheets.

If some data collectors are not subject matter experts, the analyst must insure that all of them are familiar with the tasks to be evaluated, and may have one of the data collectors who is a subject matter expert conduct this segment of the training.

In the last few steps, the analyst gives a thorough explanation and discussion of the questions on the applicable observation worksheets, demonstrates methods for interviewing trainees and administering questionnaires, and gives the data collectors an opportunity to practice making ratings on a segment of the training.

Task D2: Collect data on training and testing events. In this task, the TEE analyst sends the data collectors out to observe the training, and on a daily basis prepares assignments, stating where to go and what to observe. The analyst includes the appropriate data collection worksheets prepared in Task C2 and training materials needed for reference. When the data collectors return, the analyst reviews the data and clarifies problem areas with the data collectors, and decides whether data are usable or not and whether individual data collectors can indeed collect good data.

When a series of training events has been observed that apply to an individual task, the analyst fills out a special worksheet answering questions that are broader than can be answered by observations of a single training event.

Task D3: Collect trainee and instructor characteristics data. At some point during the TEE, the analyst examines available personnel records or other documentation to ascertain course entrance requirements and instructor qualifications, and the degree to which students and instructors possess them. The TEE Handbook provides guidelines for filling out worksheets for trainees and instructors.

Phase E: Assess Trainee Performance.

This phase has four major purposes:

a. Make assessments of how well trainees can perform the tasks selected for evaluation upon completion of training.

b. Judge the adequacy of the above assessements.

c. Summarize observations of the training and training materials as an aid to identifying performance discrepancies and as input to the revision process.

d. Identify tasks on which performance standards are not met — (performance discrepancies).

The TEE analyst takes all of the observational, interview, questionnaire, and test data collected during the TEE and synthesizes two primary outputs: a. a list of tasks for which the standards have not been met, and b. a list of potential problem areas for each task. Each deficiency will have been rated as minor or serious as it would impact test adequacy or student performance.

Phase E Tasks are:

Task E1: Collect and summarize test data. In this task, the analyst locates all of the relevant test data for the course final exam, converts to Go/No Go scoring on a task-by-task basis, if the data are not already in that form, and then enters the data on a worksheet in order to calculate the percentage of No Gos on each task and for the entire test. The same process can be accomplished for team functions, if crews rather than individuals are scored.

The data from any within-course, entry, or pretests are summarized in the same manner.

Task E2: Summarize product and process evaluation data. For process evaluation data which do not stem from master list questions, i.e., trainee and instructor reaction instruments, the reactions must be rated on a three-point scale. Having done that, the analyst averages ratings for all TEE data for questions where different data collectors have rated the same event. The averages are then used rather than the raw data. Reactions by trainees and instructors are labelled as pertaining to product or process evaluation.

Testing problems are then summarized: the analyst records the question number and rating for each problem noted in the data for each task or learning objective. Product and process ratings are recorded under separate headings, which are segmented further into sections for questions that apply to each test item, to the objective, or to the test as a whole. The analyst then uses a similar method to record ratings of the training presentation. When serious problems are identified in the next task, these summaries of training and testing problems become one of the primary TEE outputs mentioned above.

Task E3: Identify task and team function discrepancies. The analyst must first define an appropriate performance standard for each task. These standards are based on task factors rated in TEE Task A3, such as task criticality or uniqueness. The analyst then rates test adequacy for each task based on the seriousness of problem areas summarized earlier, and combines test ratings with percent No Go data for each task to yield a task rating, either acceptable, discrepant, or unknown. The analysis process is conducted to this point for both final exam and within-course test data. If both exist, the analyst combines the task ratings from each into a single set of ratings.

Having determined "combined task ratings" (or final exam task ratings alone when no within-course tests exist) from test adequacy and performance data, the analyst examines the presentation data summaries and rates the presentation for each task based on the seriousness of the problems summarized, again as acceptable, discrepant, or unknown. Finally, the analyst combines the ratings from tests and performance data with the presentation ratings for a "final task rating" for each task. These ratings represent the other primary TEE output referred to above.

If entry or pretest scores are available, the analyst can identify student selection problems by arraying entry, pretest, within-course, and final exam data for each trainee. Summary data is recorded at the bottom of the worksheet. Test adequacy is also rated. In addition, the analyst records discrepancies between specified and actual trainee characteristics.

When team functions are analyzed, and data are available for their subordinate tasks, the analyst can use a special worksheet to arrange the tasks and team functions hierarchically. He then rates the relationship between each team function and its subordinate tasks and identifies team functions which are discrepant due to team communication skills beyond individual task performances and those which are discrepant due to problems with subordinate tasks.

Phase F: Document the TEE.

The purpose of Phase F is to summarize the findings of the TEE, prepare a report documenting the effort, and organize the raw data worksheets for future use. Phase F has two tasks:

Task F1: Prepare report summarizing TEE conduct and and findings. Before writing the final report, the analyst must first consider whether all desirable data have been collected, and if not must decide whether it is feasible to collect additional data. When the data and its analysis are complete, the analyst must then interpret his findings and draw some conclusions, considering the following items:

a. Test adequacy.

b. The number of tasks rated discrepant or unknown.

c. The need for a second TEE (when the tests are very inadequate or extensive revisions in the presentation are indicated).

d. Which lessons need the most revisions.

e. Relationships between poor lessons and instructors.

f. An excessive number of No Gos on the entry test.

g. An excessive number of Gos on the pretest.

h. Relationships between trainee characteristics and entry and pretest scores.

i. Problems with team functions not related to problems with their subordinate tasks.

The analyst prepares a rough draft of his conclusions and recommendations and has some of his associates review it, taking note of their questions and challenges. The analyst prepares his complete report following sections of the TEE Handbook which give examples of techniques for displaying data and a recommended outline for the report.

Task F2: Preserve TEE documentation. So that future TEE analysts and those who may become involved in revising the course will be able to reference the TEE documentation, the analyst files it by task and step. The documentation includes course materials, product and process data collection worksheets, summary worksheets, the final report, and other TEE documentation.

SECTION VI

Discussion

Three issues encountered during the development of the TEE methodology will be discussed in this section. The first issue concerns the trade-offs that were made between the precision of the evaluation data produced and the usability of the TEE system. The second issue concerns the rationale for developing a generic evaluation system that can be tailored to almost any evaluation setting, and the third addresses the nature of TEE outputs provided for the user.

Precision vs. Usability

A number of very precise evaluation methods exists. For example, some evaluation systems look at instructor performance very closely. One such system requires data collectors to record very minute behavioral data in a multi-variable context every few seconds (Flanders, 1970). In another setting, the evaluation of training materials, it is possible to specify interval level metrics as criteria for many different questions that could be asked. For example, if one is asking whether a written test covers the course content, the evaluator can be asked to come up with the percentage of objectives covered by test items. This procedure was originally included in the TEE methodology, but was deleted in subsequent revisions due to the reasons stated below.

While evaluation techniques of the kind described above may be thought feasible and may even be applied in some academic settings, in the military setting in which TEEs must be conducted, such procedures would be much too difficult. It can also be questioned whether anything is gained by measuring many instructional variables so precisely. In the first example above, a few variations in the number of occurrences of many of the fairly minute classroom behaviors observed will not substantially affect instructional outcomes. In the second example, a difference of a few percentage points in how much the test items cover the objectives will not substantially change test content validity. Subjective measures (which involve judgments of quantity or quality, rather than precise measurements) will do just as well, and the increments of such judgments (e.g., the points on a rating scale) will be more likely to correspond to ultimate instructional outcomes.

This issue of precision versus the usability of evaluation methods was a major consideration as TEE evaluation questions and procedures were developed. These developmental criteria have been followed:

a. Evaluation questions have been constructed so as to be answered on three-point rating scales, each with rating points defined for the individual question. A concerted effort has been made to insure that the definitions of the points on each scale and the questions themselves are clear and understandable to data collectors with a potentially low reading level.

b. Several evaluation questions have been deemed too difficult for anyone without special training or a training development background to answer. Most of these questions involve the classification of content

type, a difficult procedure for anyone to learn. Such difficult questions were reserved for the TEE analyst to answer.

c. In addition to difficult questions being set aside for the TEE analyst, all questions have been constructed so that data collectors are asked only to make observations about what happened during the training, not to make judgments about how any problems observed will affect student performance. Guidance on making such judgments has been placed in a separate appendix for use by the TEE analyst during the data analysis.

d. Training procedures and instructional materials have been included for data collectors and the TEE analyst. These include instruction on key terms used in the evaluation questions, how to classify the task level of an objective, how to make ratings (for data collectors), and how to classify content type (for the analyst).

e. A few procedures that would have been too time consuming for the analyst (although not unduly difficult) have been replaced with more subjective evaluation questions. For example, one question originally included a method for ascertaining the reading grade level of instructional materials. It involved counting words, sentences, and syllables and calculating a quantitative index. This method was eliminated since it was considered too time consuming for most TEE analysts to undertake in the context of a TEE.

A Generic Evaluation System

One of the predecessor systems to the TEE methodology, Harless Performance Guild (Note 1), gave a set of example worksheets with the suggestion that these be tailored to the evaluation setting in which they were used. No further instructions were given, however, on how that should be done. Furthermore, it seemed clear that they would only apply to instructor-led training.

In order to fulfill the need for guidelines on tailoring the evaluation system to any potential training setting, a master list of evaluation questions was developed. This list contains all evaluation questions for every training setting. Short forms of these questions were distributed to worksheets for eight possible types of training events, e.g., classroom instruction, demonstration, performance test, practice, and individualized instruction. Each of these worksheets contains all master list questions that could conceivably apply to that particular setting. Procedures are given for ascertaining which questions do not apply in a given TEE and eliminating them.

The TEE methodology is thus a generic evaluation system which can be adapted by military personnel to any training situation they might encounter.

What the TEE Output Provides

The TEE results consist primarily of two sets of data:

a. An evaluation of the training for each task evaluated, either 1. acceptable, the specified number of students meet a set of valid performance standards; 2. discrepant, not enough students meet the standards; or 3. unknown, it is not known whether enough students meet the standards (probably due to inadequate tests).

b. A list of training and testing problems for each task, with each problem rated as minor or serious. A serious problem is one which would most likely cause a task to be discrepant in and of itself.

The first data set provides the user with an evaluation of the course in terms of the number of tasks which are acceptable or discrepant. Data are also available which give a rating of the adequacy of course tests. The first data set also gives the user an indication of which parts of the training are most in need of revision (i.e., those pertaining to discrepant tasks). If the tests are adequate, tasks can also be ranked for revision by their performance scores (i.e., the percent of No Gos on each task).

The second data set identifies those specific problems in the training that could cause learning problems and thus constitute possible revisions. It also identifies those problems that are serious, thus establishing a priority for revision within each task.

When data are available from entry and pretests, these data will give indications about any student selection problems the course is having. When team functions are evaluated, an analysis is available which will identify discrepancies as being related exclusively to the team function or to problems with the individual tasks of which it is composed.

Thus the TEE output provides the user with information he can use to accomplish either of two basic purposes, course evaluation or course revision.

SECTION VII

Recommendations

Validation Research

Given the sound literature base from which the TEE methodology has been derived, the systems approach used in its development, and the formative evaluation already conducted on a sample of Air Defense training, confidence in their adequacy is justified. However, to complete the TEE development process, validation research is needed. Such research would apply the TEE methodology to a complete program of instruction of Army Air Defense training (either ongoing training or training for an operational test), and would utilize Army personnel as TEE analysts and TEE data collectors.

The TEE system validation should exercise both the product and the process components of the TEE methodology on a single training package and its implementation. It is recommended that the validation be structured so as to assess the usability of the TEE documentation and the adequacy of the procedures as they are being applied, and additionally to determine the utility of the TEE outputs. Results of the validation research should document training deficiencies discovered by the TEE and the TEE system deficiencies (and proposed revisions) derived from an analysis of the TEE analyst's and TEE data collectors' working documentation and comments.

The validation research should also include an assessment of user acceptance of the TEE methodology. Care should be taken to identify the characteristics of the population which experiences the greatest success with the TEE methodology so that the most appropriate user can be targeted for future applications.

One final recommendation concerning the TEE validation research is that a resource utilization log should be maintained. The number of personnel required for data collection in a complete TEE conducted in accordance with the TEE methodology (or their predecessors) is not presently known. Neither is the requirement for the TEE analyst and associate analyst known in terms of work hours for project completion. The collection of such resource data would be invaluable to decision makers when planning future TEEs.

Follow-On Research and Development

A second area of investigation which is recommended is that which would fulfill a need previously identified by ARI and the Army system developers, and one that is complementary to the TEE methodology. This research and development requirement is to develop methodologies for applying feedback information from Air Defense TEEs to improve training packages. An important underlying assumption in instructional system development is the use of OT data initially, and quality control data from system implementation later in the development cycle for revision purposes. A failure to meet training expectations must be identified and traced back to an underlying cause before remedies can be applied. Once performance discrepancies and training deficiencies have been identified, the problem becomes one of delineating and satisfying revision requirements.

In light of the above stated needs, two specific research requirements have been specified:

a. Develop a methodology and user's guide for modifying Air Defense training packages as the result of training effectiveness evaluation; and

b. Develop recommendations for improved courseware generation procedures.

Both of these research requirements are consistent with the direction of the TEE methodology. The first requirement is an extension of the procedure for identifying training deficiencies, and the second is consistent with the ISD orientation upon which the product evaluation component of the TEE methodology has been based.

SECTION VIII

References

Reference List

Brady, R., Kincaid, J. P., & Aagard, J. A. Use of mnemonics in training materials: A guide for technical writers (TAEG Report No. 60). Orlando, FL: Training Analysis and Evaluation Group, 1978.

Branson, R. K., Rayner, G. T., Cox, J. L., Forman, J. P., King, F. J., & Hannum, W. J. Interservice procedures for instructional systems development (TRADOC Pamphlet 350-30, 5 vols.). Fort Monroe: U. S. Army Training and Doctrine Command, 1975.

Carey, J. and Briggs, L. J. Teams as designers. In L. J. Briggs (Ed.) Instructional design principles and applications (Ch. 9). Englewood Cliffs, NJ: Educational Technology Publications, 1977.

Champagne, A. B. & Klopfer, L. E. Formative evaluation in science curriculum development. Journal of Research in Science Teaching, 1974, II, 185-203.

Department of Army, TRADOC Reg 350-4, The TRADOC Training Effectiveness Analysis (TEA) System, June 1979.

Department of Army, TRADOC Training Effectiveness Analysis Handbook, First Draft. U. S. Army White Sands Missile Range, NM: U. S. Army TRADOC Systems Analysis Agency, 1979.

Dick, W. & Carey, J. The systematic design of instruction. Glenview, IL: Scott, Foresman and Company, 1978.

Ellis, J. A., Wulfeck, W. H., II. The instructional quality inventory: IV. Job performance aid. (NPRDC Special Report 79-5). San Diego, CA: Navy Personnel Research and Development Center, 1978.

Ellis, J. A., Wulfeck, W. H., II, & Fredericks, P. S. The instructional quality inventory: II. User's manual (NPRDC Special Report 79-24). San Diego, CA: Navy Personnel Research and Development Center, 1979.

Flanders, N. A. Analyzing teacher behavior. Reading, Massachusetts: Addison-Wesley, 1970.

Hanson, V. L. & Purifoy, Jr., G. R. TSM guide to training development and acquisition for major systems (ARI TR-78-A7). Valencia, Pennsylvania: Applied Science Associates, Inc. 1978. (AD A053 489)

Kern, R. P., Sticht, T. G., Welty, D., & Hauke, R. N. Guidebook for the development of Army training literature. Alexandria, VA: U. S. Army Research Institute for the Behavioral and Social Sciences, 1976 (AD A033 935).

Merrill, M. D., Reigeluth, C. M., & Faust, G. W. The instructional quality profile: A curriculum evaluation and design tool (Ch. 6). In H. F. O'Neil (Ed.) Procedures for instructional systems development. New York: Academic Press, 1979.

Schulz, R. E. & Farrel, J. R. Job aid manuals for phase II - Design of the instructional systems development model (Research Product 80-16) and Job aid manuals for phase III - Development of the instructional systems development model (Research Product 80-18). Alexandria, VA: Army Research Institute, 1980.

Shriver, E. L. Fully proceduralized job performance aids: guidance for performing behavioral analyses of tasks (AFHRL-TR-75-38). Brooks Air Force Base, TX: Air Force Human Resources Laboratory, 1975. (AD A015 059).

Swezey, R. W. & Pearlstein, R. B. Guidebook for developing criterion-referenced tests. Alexandria, VA: U. S. Army Research Institute for the Behavioral and Social Sciences, 1975.

Thurmond, P. Development of analysis, design and development techniques for team ISD. Alexandria, VA: U. S. Army Research Institute for the Behavioral and Social Sciences, 1980. (AD A090 195).

U. S. Navy. Tests, measurement of student achievement. Data Item Description Number DI-H-2033A, Author, 1 Oct. 1976.

Wagner, H., Hibbits, N., Rosenblatt, R., & Schulz, R. Team training and evaluation strategies: State-of-the-art (HumRRO-TR-77-1). Alexandria, VA: Human Resources Research Organization, 1977. (AD A027 507).

Wessling, Jr., R. G. ADA training developments. Air Defense Magazine, October-December 1979, pp. 12-19.

Reference Notes

1. Harless Performance Guild. Guidelines for conducting a training program evaluation (Working Paper FKFU 80-1). Fort Knox, KY: U. S. Army Research Institute for the Behavioral and Social Sciences Fort Knox Field Unit, 1979.

2. Kristiansen, D. M. Personal communication, 27 August 1980.

3. Roid, G., & Haladyna, T. A handbook on item writing for criterion-referenced tests (NPRDC TN 80-8). San Diego, CA: Navy Personnel Research and Development Center, 1980.

4. Courseware, Inc. Writing technically correct test items. Evaluation Workshop (Lesson 5), Courseware Instructional Design Series. San Diego, CA: Author, 1977.

5. Dick, W. Applications of formative evaluation to the instructional design process. Paper presented at the convention of the American Educational Research Association, New York, April 1977.

APPENDIX A

Abstracts Of Literature Relevant To The TEE Methodology

Introduction

This appendix contains selected abstracts from the literature of training and evaluation which were reviewed during the development of the TEE methodology. The abstracts are organized alphabetically by author in two sections, one on evaluation and training, the other on team training and evaluation. All abstracts dealing with unit, team, or collective training or evaluation are found in the latter section.

Evaluation and Training

Antonoplos, D., et al. A guide to evaluation materials, Volume II. Washington, D.C.: National Institute of Education, 1978. (ED 174 664)

Sixty-one evaluation products developed or published by universities, federal, state, intermediate, and local educational agencies, commercial publishers, private research and development agencies, and others are described to help potential users identify, select, and obtain useful materials. Potential users include teachers, administrators, community or parent groups, evaluators, program administrators, trainers, and media or curriculum specialists. The product reviews are divided into five categories: a. general training in evaluation; b. evaluation of specific subjects or kinds of education; c. evaluation of curriculum or instruction; d. needs assessment and/or goal setting; and e. evaluation validation or instrument development. Each product review contains: intended users; type of evaluation; approach to evaluation; format; product components; content emphasis; product purpose or goals; content organization or user activities; requirements; adaptability; related products; cost; availability; history and evaluation; and comments.

Ball, S., & Anderson, S.B. Practices in program evaluation: A survey and some case studies (TR-2). Arlington, VA: Office of Naval Research, 1975. (AD A017 096)

This is the second of three reports in a series of theoretical and empirical investigations of program evaluation. A questionnaire survey of 200 adult, technical training programs and their evaluations was conducted. This was followed up by site visits to 14 of them for in-depth study. The 200 programs were divided equally among Department of Defense, other federal government departments and agencies, state

and local governments and agencies, and private sector commercial, business, and industrial organizations. The survey was useful in isolating areas of concern for further investigation.

Beason, G.M. Bias in performance evaluation: An examination of the relationship of the rater to the ratee (AFOSR TR-77-0988; WSUCHA TR-116). Wichita, KS: Wichita State University, Center for Human Appraisal, 1977. (AD A043 230)

The effects of personality variables on rating behavior were studied. The experienced rater showed more bias than non-raters and it was related to their own personality characteristics and role preferences. Other bias was found favoring extroverts and derogating independents and neurotics. Rating factors were identified.

Bell, N.T., & Abedor, A.J. Developing audio-visual instructional modules for vocational and technical training. Englewood Cliffs, NJ: Educational Technology Publications, 1977.

This book contains three major elements. First, it presents a validated sequence of steps to be followed in developing audio-visual modules for vocationally related instruction. Second, the book describes a set of critical elements of instruction which are ignored by most designers of instructional materials. Third, the book presents a new, but potent instructional strategy, which takes account of the critical elements of instruction as they specifically apply to vocationally related learning.

Bergman, B.A., & Siegal, A.I. Training evaluation and student achievement measurement: A review of the literature (AFHRL TR-72-3). Wayne, PA: Applied Psychological Services, 1972. (AD 747 040)

The current training evaluation and student measurement literature is reviewed. The emphasis is on studies which have been reported in the last ten years, although earlier studies which have impacted heavily on recent trends are also included. Because of the obvious interaction between both training evaluation and student measurement, on the one hand, and such topics as statistical methods, methods for course development, training methods, learning styles, motivation, and moderator variables, on the other hand, these and similar considerations are also included.

Bond, N.A., Jr., & Rigney, J.W. Measurement of training outcomes (TR-66). Los Angeles: University of Southern California, Department of Psychology, 1970. (AD 711 302)

Measurement of training outcomes as a requirement for evaluating new training techniques is one that is difficult to meet. Managers may have different goals from those of the investigators. In the report, possibilities for measuring outcomes of training are surveyed, viewing training as a form of planned social change. Approaches which are discussed include adaptive control models, decision theory models, and simulation models. Illustrations from the computer assisted instruction of recent attempts to measure training outcomes are given.

Borich, G.D. (Ed.). Evaluating educational programs and products. Englewood Cliffs, NJ: Educational Technology Publications, 1978.

This book is a guide and handbook for planners, developers, and evaluators of educational programs and products. It provides practical insights that are immediately applicable to planning and executing effective program and product evaluations.

The book divides the evaluator's work into three important activities: establishing perspective, planning the evaluation, and analyzing the data. The first activity is completed when the evaluator chooses an appropriate role for the context in which he will work; the second when he chooses an appropriate model or strategy for planning the evaluation; and the third when he selects appropriate methods and techniques for analyzing the data. The key to each of these activities is the word "appropriate." The task of this book is to identify specific procedures that are appropriate to each of these activities.

Borich, G.D. A systems approach to the evaluation of training. In H. F. O'Neil, Jr. (Ed.), Procedures for instructional systems development (Ch. 7). New York: Academic Press, 1979.

This chapter introduces a specific systems approach for conducting evaluations of training, presents a general model for the evaluation of training that incorporates this approach, and illustrates how various stages of the model can be employed to improve the structure and content of a training program. The overall objective of this chapter is to provide a coherent, integrated systems approach to planning, developing and evaluating training programs.

Braby, R., Kincaid, J.P., & Aagard, J.A. The use of mnemonics in training materials: A guide for technical writers (TAEG Report No. 60). Orlando, FL: Training Analysis and Evaluation Group, 1978.

This report is a guide for incorporating different mnemonic techniques into the training curriculum. It is intended mainly for those responsible for the production of written training materials; classroom

instructors should also find it useful. The Interservice Procedures for Instructional Systems Development call for the use of mnemonics in curriculum development. This report provides guidelines for choosing when to use mnemonics, which types of mnemonics to use, and how to develop each type of mnemonic. It contains a description of nine techniques including several first letter mnemonics, rhymes as mnemonics, patterns and graphics as mnemonics and such special techniques as stories and the peg word method. It is filled with examples pertinent to Navy training including three complete sets of mnemonics for the teaching of: a. morse code, b. signal flags, and c. orders to the sentries.

Branson, R.K., Rayner, G.T., Cox, J.L., Furman, J.P., King, F.J., & Hannum, W.J. Interservice procedures for instructional systems development (TRADOC Pamphlet 350-30, 5 vols.). Fort Monroe, VA: U.S. Army Training and Doctrine Command, 1975.

The ISD model consists of procedures grouped into five phases:

> In Phase I, Analyze, inputs, processes, and outputs are all based on job information. An inventory of job tasks is compiled and divided into two groups: tasks not selected for instruction; and tasks selected for instruction. Performance standards for tasks selected for instruction are determined by interview or observation at job sites and verified by subject matter experts. The analysis of existing course documentation is done to determine if all or portions of the analysis phase and other phases have already been done by someone else following the ISD guidelines. As a final analysis phase step, the list of tasks selected for instruction is analyzed for the most suitable instructional setting for each task.
>
> Beginning with Phase II, Design, the ISD model is concerned with designing instruction using the job analysis information from Phase I. The first step is the conversion of each task selected for training into a terminal learning objective. Each terminal learning objective is then analyzed to determine learning objectives and learning steps necessary for mastery of the terminal learning objective. Tests are designed to match the learning objectives. A sample of students is tested to ensure that their entry behaviors match the level of learning analysis. Finally, a sequence of instruction is designed for the learning objectives.
>
> Phase III, Develop, begins with the classification of learning objectives by learning category so as to identify learning guidelines necessary for optimum learning to take place. Determining how instruction is to be packaged and presented to the student is accomplished through a media selection process which takes into account such factors as learning category and guideline, media

characteristics, training setting criteria, and costs. Instructional management plans are developed to allocate and manage all resources for conducting instruction. Instruction materials are selected or developed and tried out. When materials have been validated on the basis of empirical data obtained from groups of typical students, the course is ready for implementation.

In Phase IV, Implement, staff training is required for the implementation of the instructional management plan and the instruction. Some key personnel must be trained to be managers in the specified management plan. The instructional staff must be trained to conduct the instruction and collect evaluative data on all of the instructional components. At the completion of each instructional cycle, management staff should be able to use the collected information to improve the instructional system.

In Phase V, Control, evaluation and revision of instruction are carried out by personnel who preferably are neither the instructional designers nor the managers of the course under study. The first activity (internal evaluation) is the analysis of learner performance in the course to determine instances of deficient or irrelevant instruction. The evaluation team then suggests solutions for the problems. In the external evaluation, personnel assess job task performance on the job to determine the actual performance of course graduates and other job incumbents. All collected data, internal and external, can be used as quality control on instruction and as input to any phase of the system for revision.

Carey, J., & Briggs, L.J. Teams as designers. In L. J. Briggs (Ed.), Instructional design principles and applications (Ch. 9). Englewood Cliffs, NJ: Educational Technology Publications, 1977.

This chapter discusses the instructional design steps a team of designers would take following the completion of steps common to both an individual teacher and a design team who are developing instruction. The remaining steps are:

a. Select the type of stimulus for each instructional event for each enabling objective.

b. Select the media for each such event.

c. Select the desired conditions of learning by which each event is to achieve its purpose.

d. Write prescriptions for how the conditions of learning

are to be incorporated into each event; these prescriptions aid the media production specialists in both the content to be presented and how it is presented in each selected medium.

e. Develop and produce the instructional materials and the associated learner activity guides and tests over the objectives.

f. Conduct formative evaluation to improve the items listed in item e above.

g. Assist teachers in the use of the complete instructional system.

h. Assist teachers by monitoring the use of the system to see that all intended products and processes are being used as intended.

i. Assist with field tests, and eventually with summative evaluation of the system.

j. Assist with diffusion efforts when the system is intended for widespread application.

Carey, J.O., & Carey, L.M. Using formative evaluation for the selection of instructional materials. Journal of Instructional Development, 1980, 3, 12-18.

Instructional materials selection practices vary widely in the way they are administered and conducted, the criteria that are used, and the precision with which they are carried out. In this paper a two-phase instructional materials selection process is presented. The process is based on considerations from the design and formative evaluation of competency-based instruction. The purpose of the first phase is to select materials that have the best potential for affecting learning outcomes desired by a local or state educational agency. The purpose of the second phase is to verify decisions made in phase I and make recommendations to teachers about how the materials can be used most effectively. The paper also includes a comparison between the guidelines for materials selection published by the State of Florida and the considerations recommended in this paper.

Champagne, A.B., & Klopfer, L.E. Formative evaluation in science curriculum development. Journal of Research in Science Teaching, 1974, II, 185-203.

In this article, a lengthy set of evaluation questions and a method for the formative evaluation of science curricula are described. The questions are in six sets dealing with a. conceptualization and planning of the curriculum, b. the quality of student instructional materials and scientific apparatuses, c. short and intermediate range student behaviors (dependent variables), d. classroom management considerations, e. the functioning of teacher materials, the teacher in the classroom, and teacher preparation, and f. the marketability of the curriculum.

The formative evaluation methodology consists of four stages which employ the six sets of questions described. The first stage (A) involves subject matter expert review of the program's conceptualization and planning and a critical review of student materials. Stage B is a small-scale observation of the interactions of students with instructional materials concentrating on short-term students behaviors. Stage C is an expansion of Stage B to several classes in several schools and concentrates on both short-term and intermediate-range behaviors. Stage D concentrates on teacher preparation and the ease of program implementation with a still larger sample of classrooms. Marketability is also assessed by observing the rate of purchase by schools and use after one year's implementation.

Cole, H.R. Evaluative indices for curriculum materials and educational programs. Washington, D.C.: Bureau of Educational Personnel Development, 1975. (ED 128 319)

This training package of evaluative indices for process curriculum materials and educational programs is composed of ten handouts: a. a set of materials designed for use by teachers, curriculum coordinators, school administrators, college professors, or educational consultants, intended to teach basic concepts about process education and demonstrate how the basic objectives of any curriculum innovation may be translated into a set of indices useful for operationalizing and evaluating the program; b. a discussion of (1) assumptions, justifications, and definitions for process education, (2) opposed value positions underlying process and conventional educational practice, (3) the relationship between basic value positions and operational classroom role descriptions, and (4) translating role descriptions into appropriate and inappropriate behavioral indices for teachers and pupils; c. presentation of pupil and teacher role indices, each related to one or more of the basic value positions for process education and the derivative role expectations used to assess the degree to which the teacher and pupils in a given classroom are exhibiting behavior consistent with the goals of process education; d. an experience in creating evaluative indices; e. presentation of a curriculum in social interaction, self-perception skills, and creative thinking and feeling skills; f. an actual problem concerning an introductory teacher education program presented as a case study with questions and a set of solutions;

g. another case study problem; h. case study in goals, rationales, and procedures; i. a case study in operationalizing plans and objectives intended as a further illustration of how the general principles outlined in the first portion of the training package can be applied to teacher education program development; and j. concluding remarks.

Courseware, Inc. Writing technically correct test items. Evaluation Workshop (Lesson 5), Courseware Instructional Design Series. San Diego, CA: Author, 1977.

This lesson contains instructional materials for use in a workshop. Each lesson segment contains statements, examples, and practice on writing test items. The lesson covers the construction of technically correct performance, true-false, multiple-choice, matching, fill-in-the-blank, short answer, and listing items which are consistent with the level of task and content specified in a given objective.

Davies, J.E. A plan for the evaluation of leadership training in the United States Army (Master's thesis). Monterey, CA: Naval Postgraduate School, 1980. (AD A091 094)

The Army, in a period of constrained resources and increasing demands on its leaders, can ill afford to pursue leadership training which is ineffective. The evaluation plan developed in this study seeks to provide the decision maker with information necessary to guide the training development towards its desired outcome: producing better leaders. A review of the leadership theories contributing to the Army's organizational leadership model, their training programs, and the leadership training of the other services is presented. Their methods of program evaluation are studied. The evaluation plan is a systematic study employing five principal criteria: process evaluation, learning, attitudinal change, behavioral change, and the change in organizational performance. Each is discussed in contribution to the overall understanding of the training program's effectiveness. The evaluation scheme is presented in an action plan format to coincide with other ongoing initiatives in the leadership and educational fields.

Diamond, R.M., & Sudweeks, R.R. A comprehensive approach to course evaluation. Journal of Instructional Development, 1980, 4, 28-34.

Evaluation is an important phase of course development and improvement efforts. This article discusses a number of problems with current approaches to course evaluation. A broader, more comprehensive approach is recommended, and a checklist illustrating the kinds of issues and questions that need to be considered is presented.

Dick, W. Evaluating programmatic impact in education. Washington, D. C.: Office of Education, Teacher Corps, 1976. (ED 132 135)

This document reports the first-year activities of Teacher Corps projects demonstrating the training framework entitled Adaptation of Research Findings. These projects incorporate into their design the results of research, empirical practices, and processes that have proven effective and relevant to the educational processes for schools serving low-income populations. Chapter I provides an overview. Chapter II presents an in-depth look at the evaluation process and discusses the differences between impact and process evaluations. A comparison is made among projects that focus on either student outcomes, teacher outcomes, or institutional outcomes, and the implications of these different foci are discussed. Considerations of where to begin to conceptualize the evaluation process are presented in Chapter III. The fourth chapter discusses the design of impact evaluation studies and presents some alternative approaches to evaluation such as quasi-experimental designs and the establishment of criterion standards. Chapter V discusses the design and selection of evaluation instruments. A variety of instruments are considered as well as behavioral indicators that can be used to evaluate project outcomes. Chapter VI argues that the process of evaluating ongoing activities during the course of the project is of critical importance both to the management of the project and to the eventual sharing of the project's outcomes. The importance of careful preparation of data gathered during the evaluation process is considered in Chapter VII, and the importance of the organization, display, and interpretation of data in order to maximize usefulness is emphasized. Chapter VIII focuses on some of the major problems that arise in the impact evaluation process. The final chapter summarizes the importance of both process and impact evaluations.

Dick, W. Applications of formative evaluation to the instructional design process. Paper presented at the convention of the American Educational Research Association, New York, 1977.

Formative evaluation is normally applied to prototype instructional materials prior to their final production. The application of such procedures at earlier stages of the instructional design process, however, may be of great benefit. General procedures for the formative evaluation of six steps in the ISD process are proposed. Examples detailing two of these procedures are given: formative evaluation of learning hierarchies and of student entry behaviors and characteristics. Both procedures involve the use of test items (which can also be formatively evaluated at the same time) to collect empirical data about students' abilities.

Dick, W., & Carey, J. The systematic design of instruction. Glenview, IL: Scott, Foresman and Company, 1978.

After discussing the origins of systematically designed instruction, a model for the systematic design of instruction is presented. It has ten components: a. identifying an instructional goal, b. conducting an instructional analysis, c. identifying entry behaviors and characteristics, d. writing performance objectives, e. developing criterion-referenced tests, f. developing an instructional strategy, g. developing instructional materials, h. designing and conducting formative evaluations, i. revising instructional materials, and j. summative evaluation and grading.

Dieterly, P.L. The evaluation of training with specific emphasis on criteria (AU AFIT SL-9-73). Wright-Patterson Air Force Base, OH: Air Force Institute of Technology, School of Systems and Logistics, 1973. (AD 771 009)

A review of the literature on training evaluation is presented. An attempt is made to establish a perspective of the current status of evaluating training programs that occur in the industrial, military, educational, and governmental systems. Emphasis is placed upon the traditional problem of criterion measurement and a suggested model is introduced for evaluating a major training program. The paper provides a comprehensive introduction into the problems of training evaluation.

Downey, R.G., & Duffy, P.J. Review of peer evaluation research (ARI TP-342). Alexandria, VA: U. S. Army Research Institute for the Behavioral and Social Sciences, 1978. (AD A061 780)

Peer evaluation research was reviewed from the three major perspectives of validity studies, methodology, and situational factors. Most of the research programs were conducted in the course of developing procedures for evaluating training groups (e.g., in Officer Candidate School, U. S. Military Academy, and Ranger course). Substantial concurrent and predictive validity generally was found, with correlation coefficients in the .30 to .50 range. Different evaluation methods (rating, ranking, nominations, and combinations of these techniques) did not differ substantially in either reliability or validity. Evaluation methods did, however, vary in acceptability and feasibility. A review of the documented and potential effects of situational factors impacting on the evaluation process indicated that users of peer evaluation should be aware of these issues in designing programs. Many issues surrounding peer evaluations remain unresolved. Evidence suggests that these issues can be resolved, and that they do not detract from the conclusion that peer evaluations are a powerful tool in discriminating complex human behavior.

Dyer, F.N., et al. A method for obtaining post formal training feedback: Development and validation. Final Report (TAEG R-19). Orlando, FL: Naval Training Equipment Center, Training Analysis and Evaluation Group, 1975. (ED 110 032)

An evaluation of alternative methods--including three types of questionnaires and face-to-face interviews--of obtaining post training feedback from naval personnel is described. Using a sample of recent radio technician trainees, various approaches to data collection were employed. It was found that questionnaires provide the most cost-effective means of obtaining the needed information. Recommendations for further development and implementation of the questionnaire procedures are presented. The appendixes include the data collection instruments used, interview instructions, and summary data sheets.

Elsbree, A.R., & Howe, C. An evaluation of training in three acts. Training and Development Journal, 1977: Act I - Focus, July, pp. 10-14; Act II - Plan, August, pp. 12-19; Act III - Implement, September, pp. 20-35.

This series of articles portrays a three-part process for the evaluation of training programs. The process is a sequence of activities and decisions geared toward making evaluations responsive to the information needs of clients, i.e., people making decisions about training.

Each article covers one of the three phases of the process. In the Focus phase, the evaluator establishes the extent and objectives of the evaluation effort. The blueprint for conducting the project is produced in the Plan phase. During the Implementation phase, the evaluator puts the plan into action to obtain necessary data, interpret them, and provide information to the client.

These articles do not attempt to discuss each activity and decision point of the evaluation process in detail, but rather present a dramatization with commentary. The articles walk through a simulated evaluation of a training program to enable the reader to envision how the process might be applied in a "real" situation.

Since the choice of methodology, design, and instrumentation depends upon the specific situation, the series is necessarily confined to the particulars of the example.

Gooler, D.D. Formative evaluation strategies for major instructional development projects. Journal of Instructional Development, 1980, 3, 7-11.

The paper contains three major sections. First, an attempt is made to identify some general maxims that seem to shape major instructional development efforts and thus affect efforts to do formative evaluation within those projects. In the second section, four major issues are described that must be addressed in any attempt to plan for and implement formative evaluation as part of a large development effort. The third section focuses on some simple but practical procedures that might enhance the probability of formative evaluation being effectively used as a part of a major instructional development effort. These procedures, taken together, constitute a strategy.

Gropper, G.L. Diagnosis and revision in the development of instructional materials. Englewood Cliffs, NJ: Educational Technology Publications, 1975.

This is a state-of-the-art volume on the diagnosis and revision of instructional materials. It neither discusses procedural models described in the literature nor does it describe one of its own. It does provide a description of a variety of diagnostic issues and methods which a student, a teacher, a developer, or a researcher may find useful when reading other accounts of the tryout and revision process. It also provides a description of a range of tools and methods for diagnosis and revision from which the reader can select and use in any combination those most appropriate to his own current research or development needs.

Hall, E.R., et al. A comparative assessment of three methods of collecting training feedback information. Final report (TAEG R-64). Orlando, FL: Naval Training Equipment Center, Training Analysis and Evaluation Group, 1978. (ED 174 660)

Three methods of obtaining training feedback data from recent Atlantic Fleet technician school graduates and their fleet supervisors were compared: a. a mailout questionnaire; b. a structured interview; and c. a job knowledge test. The results demonstrated that the questionnaire and structured interview procedure produced equivalent rating scale data concerning adequacy of training, frequency of task performance, and supervisors' assessments of graduate proficiency. Ratings of training adequacy and frequency of task performance obtained from school graduates were equivalent to those obtained from graduates' supervisors. In the short run, questionnaires were least expensive for data collection, and job knowledge tests were the most expensive. Over the long term, with larger populations of graduates, job knowledge tests became less expensive, while structured interview costs remained high. Selection of methods for data collection must also consider the specific information needs to be met, plus the relative power of each method for producing the needed information.

Hall, E.R., Lam, K., & Bellomy, S.G. Training effectiveness assessment: Volume I, current military training evaluation programs. Final report (TAEG R-39). Orlando, FL: Naval Training Equipment Center, Training Analysis and Evaluation Group, 1976. (ED 137 390)

A study was conducted to clarify issues and problems involved in the assessment of the effectiveness of military training and to evaluate and recommend objective procedures for determining the effectiveness of Navy training. The study results are reported in two volumes. This volume reviews current military training evaluation programs. Evaluation philosophy, documentation, and current practices in the assessment of training effectiveness within the United States Air Force, Navy, Marine Corps, and Army are described. Information is provided concerning strengths and apparent deficiencies of the programs which were in effect between June 1975 and May 1976.

Hall, E.R., Rankin, W.C., & Aagard, J.A. Training effectiveness assessment: Volume II. Problems, concepts, and evaluation alternatives (TAEG Report No. 39, Volume 2). Orlando, FL: Training Analysis and Evaluation Group, 1976. (AD A036 518)

A study was conducted to clarify issues and problems involved in the assessment of the effectiveness of military training and to evaluate and recommend more objective procedures for determining the effectiveness of Navy training. The study results are reported in two volumes. Volume II examines specific problems affecting Navy training evaluation programs. It provides discussions of technical considerations relevant to the conduct of evaluation and training effectiveness assessment. General procedures for assessing the effectiveness of Navy training courses are given, and a number of methodological options for evaluation data gathering are described and evaluated.

Hanson, R., et al. The development of classroom observation procedures for evaluating training (SWRL TM-[illegible]-71-07). Los Alamitos, CA: Southwest Regional Laboratory for Educational Research and Development, 1971. (ED 110 516)

This report describes the procedures followed in developing classroom observation procedures for use in evaluating the First Year Communication Skills Program and Instructional Concepts Program training systems. The procedures cover the identification of the variables to be measured, development of scales to measure these variables, and the training of observers in the use of these scales.

Harless Performance Guild. Guidelines for conducting a training program evaluation (Working Paper FKFU 80-1.) Fort Knox, KY: U. S. Army Research Institute for the Behavioral and Social Sciences, Fort Knox Field Unit, 1979.

This document is a set of guidelines and data collection instruments for conducting a Training Program Evaluation. They are intended to be applicable to any Army training, but are based on a review of training and documentation concerning the XM-1 tank. Each task described in the guidelines and its accompanying worksheet are for the purpose of either collecting data, summarizing data, or interpreting data. The Guidelines give step-by-step directions and examples for using the worksheets to conduct an evaluation.

The approach taken for an evaluation is built around five "phases":

a. Phase A: Plan the TEA
b. Phase B: Observe training and testing
c. Phase C: Assess quality of trainee performance
d. Phase D: Hypothesize and investigate training-causes of deficiencies
e. Phase E: Document findings of the TEA

The scope of an evaluation project includes collecting data for the purposes of identifying "failures" in performance of tasks by the trainees (called "performance deficiencies" in the Guidelines). The scope also includes methods for determining if these deficiencies were probably caused by a deficiency in training received, rather than in training management considerations or selection problems. The scope of the project does not include "how-to-fix" any training-caused deficiencies in performance.

Jeantheau, G.G. Handbook for training systems evaluation (DAC-69-129). Orlando, FL: Naval Training Device Center, 1971. (AD 885 751)

The handbook presents the procedures for conducting evaluations of the effectiveness of training in training devices. Four levels of evaluation are treated: qualitative assessment, non-comparative measurement, comparative measurement, and transfer of training. Each succeeding level provides increasing rigor but also entails increased problems of coordination and cooperation with the training activity. A field trial of the method with Device 21A39, Submarine Attack Teacher, is discussed and examples of materials and procedures are given. Recommendations are included for application of the method to other training device settings.

The Joint Committee on Standards for Educational Evaluation. Standards for evaluations of educational programs, projects, and materials. New York: McGraw-Hill Book Company, 1981.

This book is for the use of persons who commission, conduct, or employ the results of evaluations to improve education: teachers, administrators, evaluators, curriculum specialists, school board members, legislators, counselors, leaders of educational associations, parents, and others. It is for those who work in or are concerned about elementary, secondary, higher, or adult education, and it is intended for use in both private and public institutions. It is a guide to be used in evaluating educational programs, projects, and materials.

The book identifies and elucidates 30 separate standards. They are presented in four groups that correspond to four main concerns about any evaluation--its utility, feasibility, propriety, and accuracy. Each standard is explained and clarified through a commentary which includes an overview of intent, guidelines for application, common pitfalls, caveats (or warnings against being overzealous in implementing the standard), and an illustration of the standard's application.

Jorgensen, C.C. Early training assessment within developing system concepts (ARI RR-1224). Alexandria, VA: U. S. Army Research Institute for the Behavioral and Social Sciences, 1979. (AD A082 916)

This paper presents a proposal for training assessment within early system concepts. A broad spectrum of training requirements generated by recent Army guidance for determining training impacts at the earliest stages of weapon system specification is considered. An examination of the state-of-the-art is made along with recommendations for six methodological areas: concept generation, task specification, trade-off analysis, management information, system effectiveness estimation, and costing. Innovative and little known techniques discussed include both tri-service and foreign research. A proposal is made for combinations and extensions of existing research to meet projected Army needs. Areas in need of further research are identified.

Kern, R.P., Sticht, T.G., Welty, D., & Hauke, R.N. Guidebook for the development of Army training literature. Alexandria, VA: U. S. Army Research Institute for the Behavioral and Social Sciences, 1976. (AD A033 935)

The guidebook is a complete job aid for the writer of performance oriented training literature, with step-by-step instruction for the

development of operators' and maintenance manuals and special training/performance texts at the appropriate reading level. Numerous examples of proper text preparation, best use of graphic illustration, selected reading levels, and practical motivation of the user of the training/performance literature are presented.

Knerr, C.S., Barton, H.D., Lombardo, J.F., Sr., & Katz, M.S. Evaluation instruments for the basic noncommissioned officer course for combat arms soldiers (ARI Res Problems Rev-77-9). Alexandria, VA: U. S. Army Research Institute for the Behavioral and Social Sciences, 1978. (AD A076 700)

An evaluation program and questionnaire was developed for a pilot basic noncommissioned officer course in the combat arms (BNCOC/CA). ARI designed two types of instruments for evaluating and refining the pilot course; a. questionnaires to obtain subjective judgements from students and instructors and b. task-performance score forms to record MOS-specific training data. Instruments were tailored to 10 different MOSs. The complete package consisted of a manual for administering the evaluation program, performance-based tests of skills covered in courses, and summary reports of validity of tests and of their application. A model and method for achieving quality control in lower and medium NCO courses was also designed. Not only did TRADOC and the service schools use the instruments to revise and judge the effectiveness of the pilot course, but the score forms serve as an operational training record for the NCO Academy, for individuals and entire classes. The BNCOC/CA course was implemented worldwide in 1977, with an expected 10,000 graduates each year. Its successful evaluation and implementation are a major contribution to the enlisted personnel management system.

Law, A.I., & Bronson, W.H. Program evaluator's guide. Washington, D.C.: Office of Education, 1977. (ED 142 563)

This guide presents detailed information concerning the purposes and process of program evaluation, the role of the evaluator, and the development of an evaluation plan or design. Instruction is provided in selecting or developing assessment instruments, collecting and analyzing data, reporting evaluation results, and applying the findings. The manual, which includes learning exercises, was developed under the California Evaluation Improvement Project as a study guide for use in inservice training workshops for program evaluators, teachers, principals, curriculum specialists, and other individuals responsible for school programs and those who aid educational administrators in ascertaining program effectiveness.

Melching, W.H., Brennan, M.F., Hungerland, J.E., Showel, M., & Taylor, J.E. The development and trial evaluation of alternate programs for unit training managers and trainers (HumRRO FR-WO-CA-75-23). Alexandria, VA: Human Resources Research Organization, 1977. (AD A042 586)

The goal of this effort was to develop alternate training programs to enable unit training managers and trainers to employ performance-based practices in training and in evaluating individuals in their performance in the unit. The work was performed in conjunction with the 7th Infantry Division, Fort Ord, California. Using official Army guidance documents, training manager and trainer functions were analyzed into tasks. Performance objectives were then developed and used to guide construction of performance tests and training programs. Two training approaches were employed. Directed practice (DP) and guided self study (GSS) programs were developed for both manager and trainer. The DP programs involved frequent face-to-face interactions between managers/trainers and training experts, and gave the student opportunity to practice the desired training skills. The GSS programs relied mainly on specially prepared guidance materials, and they also gave the manager and trainer frequent opportunity to practice the desired skills. All programs were self-contained. Preliminary versions of the programs were administered to 98 unit managers. Final programs were tested on 19 division personnel. Limited evaluation indicated the programs were effective and of utility to division personnel.

Merrill, M.D., Reigeluth, C.M., & Faust, G.W. The Instructional Quality Profile: A curriculum evaluation and design tool. In H. F. O'Neil (Ed.) Procedures for instructional systems development (Ch.6). New York: Academic Press, 1979.

The Instructional Quality Profile (IQP) provides a set of detailed procedures for analyzing the quality of instruction in relation to different kinds of objectives and test items. Instructional quality refers to the degree to which instruction is effective, efficient, and appealing--that is, the degree to which it works in cost-effectively promoting student performance on a posttest and student affect toward learning. Educators have developed detailed procedures for making reliable tests and for writing well-stated objectives. However, very little attention has been devoted to detailed procedures for analyzing instruction. The Instructional Quality Profile is an analytic tool for diagnosing specific weaknesses and correcting those weaknesses in existing instruction and for providing prescriptions for avoiding such weaknesses in the design of new instruction.

The Instructional Quality Profile has also been referred to as the Instructional Strategy Diagnostic Profile and the Instructional Quality

Inventory. This chapter gives a thorough description of the framework of the IQP, where it fits in the context of a total performance system, and how it is applied. Its potential applications include diagnosing the weaknesses of instruction, rating instruction, revising instruction, designing new instruction, and prescribing effective study skills. Research supporting the IQP is also discussed.

Military Testing Association. Proceedings of the annual conference of the Military Testing Association (18th). Pensacola, FL: Naval Education and Training Program Development Center, 1976. (ED 147 344)

The 75 papers included in these conference proceedings discuss testing conducted by the different branches of the armed forces. The importance of relating necessary job skills to the skills measured by the tests administered to the job applicants is emphasized. Various evaluation methods--including peer rating, aptitude testing, adaptive testing, performance or skill qualification testing, computer assisted testing, and job knowledge analysis--are used for personnel selection and evaluation regarding advancement. Additional topics discussed at the symposium included: the evaluation of military training programs, job satisfaction surveys, impact of female personnel in the military, and test construction. The by-laws of the Military Testing Association are appended.

Olmstead, J.A., et al. Research on utilization of assessment results and methods. Final technical report (HumRRO FTR-D4-74-18). Alexandria, VA: Human Resources Research Organization, 1974. (ED 128 439)

The Army has established an Assessment Center Pilot Program at The Infantry School, Fort Benning, Georgia. The purpose of the program is to determine the feasibility of assessment centers for the Army. The project described in this report was designed to contribute to two particular objectives of the pilot program. These objectives were a. to identify potential uses of assessment results and techniques in accomplishment of the leadership development mission of the Infantry School and the Army, and b. to develop ways of improving assessment procedures and methodology for use by the Army. To fulfill these objectives, four discriminable tasks were undertaken: a. to investigate potential uses of assessment results; b. to identify potential uses of assessment methods for training; c. to develop procedures for training assessors to use observational and recording techniques; and d. to develop a model for designing assessment exercises, or situational tests.

Osborn, W.C. Process versus product measures in performance testing (HumRRO PP-16-74). Alexandria, VA: Human Resources Research Organization, 1973. (ED 102 206)

Performance tests are used in training evaluation to a. certify student achievement, and b. diagnose weaknesses in the instructional system. Proficiency measures that focus on task outcomes (product) normally provide data relevant to the first purpose, whereas measures of how the tasks are carried out (process) pertain to the second. Time or cost factors sometimes preclude the use of product measures, leaving measures of task process as the only available criteria for evaluating training outcomes. Instances in which process measures are typically substituted for product measures are described in this paper with reference to the types of tasks for which the substitution is valid and those for which it is invalid. Theoretical and practical issues pertaining to the use and misuse of process measures are discussed.

Osborn, W.C., Ford, S.P., Moon, H.L., Root, R.T., & Word, L.E. Development of new training concepts and procedures for unit trainers (HumRRO FR-CD(L)-75-3A). Fort Monroe, VA: U. S. Army Training and Doctrine Command, 1976. (AD A024 207)

This report describes the development and testing of a 10-hour course of instruction designed to teach officers and NCO's how to manage and conduct performance-oriented training in their units. The first three hours of the course present the principles and techniques of effective performance-oriented training; the remaining lessons present practical exercises, done in small groups. The UTRAIN course has been implemented in the infantry officer basic course at Fort Benning, GA, and adapted for NCO, school faculty, and National Guard courses.

Pritchard, D., et al. Incentive motivation techniques evaluation in Air Force technical training (AFHRL TR-74-24). Brooks Air Force Base, TX: Air Force Human Resources Laboratory, 1974. (ED 106 467)

The report describes an 18-month research project at Chanute Air Force Base, Illinois, designed to evaluate the effectiveness of incentive motivation techniques in Air Force technical training. The first phase of the research identified incentives. The findings were used in the second phase of the research which made these incentives contingent on performance in two of the resident training courses at the base. The first system gave performance based incentives in the courses. The second utilized a system that attempted to give effort based incentives, while the third used financially based incentives. Research results indicated that while secondary performance measures such as amount of remedial instruction, frequency of probations, and frequency of course failures decreased under the incentive program, the primary performance

measures of exam scores and speed of course completion did not generally show much improvement. Yet, from a cost-effectiveness viewpoint, even the relatively small (i.e., 8 percent) increase in speed of course completion was meaningful. Attitudes to the program generally improved or stayed the same. The financially based incentive system was found to be the most cost-effective for Air Force technical training. A 150-page appendix provides background information, incentive attractiveness data, questionnaires, manuals, and item statistics.

Provus, M. Discrepancy evaluation for educational program improvement and assessment. Berkeley, CA: McCutchan Publishing Corporation, 1971.

This book includes the critical contributions of well-known evaluation specialists, such as John Goodlad, Egan Guba, and Dan Stufflebeam, and directors of educational research across the country. It is a practitioner's guide to the evaluation of public school programs based on the concept that evaluation is the art of describing a discrepancy between expectation and performance. The discrepancy evaluation model presented here, under formal development for five years, is applied to school district programs, state agencies, and federal educational programs.

The author argues that his discrepancy evaluation model, which includes both the case study method and experimental design (as well as other disciplined techniques), is required to conduct meaningful evaluations.

Rayner, G.T. An empirical study of a methodology for the revision of systematically designed educational materials. Washington, D. C.: Office of Naval Research, Personnel and Training Research Programs Office, 1972. (ED 067 877)

A project was devised to develop and test a revision model for systematically designed educational materials based on the literature and on previous procedures. The model divides the revision process into content changes and procedural changes, and decisions are based on data collected from measures of student performance and attitudes and on judgements of a content expert and educational technologist. Students in a required course in health education were subjects for the implementation phase of the project; they were randomly assigned to either computer managed instruction (CMI) or to conventional instruction. The implementation of the revision model was based on the results from the first quarter of operation, where only 17 percent of the CMI students reached criterion of 80 percent on the final, criterion-referenced examination. After revision of the course according to the model, 71 percent of the students in the CMI group reached criterion. While the model was clearly successful, several revisions could be made concerning criterion measures, data collection procedures, evaluation instruments, and student pacing.

Rice, D., et al. Educational evaluators--A model for task oriented position development. Contemporary Education, 1970, 41, 115-118. (EJ 015 769)

An outline of 44 evaluator tasks is discussed in terms of its usefulness in defining, evaluating, and improving the position of the educational evaluator; in adapting the position to the needs of particular institutions; and in designing appropriate evaluator training programs.

Roid, G., & Haladyna, T. A handbook on item writing for criterion-referenced tests (NPRDC TN 80-8). San Diego, CA: Navy Personnel Research and Development Center, 1980. (Preliminary document, limited distribution.)

This handbook provides a simple method for test item construction and some practical guidelines on item writing. The method consists of four basic steps. First, classify the learning objective using the system in the Instructional Quality Inventory and identify the appropriate item format. Second, follow the practical guidelines provided in this handbook for the drafting of items. Third, use the IQI to review the drafted items for consistency in matching the learning objectives. Fourth, administer the items to students to detect item flaws and then review the items accordingly.

The practical guidelines given in the handbook include directions on the actual wording and form of items, as well as rules for writing recognition, recall, and performance test items.

Schulz, R.E., & Farrel, J.R. Job aid manuals for phase II - Design of the instructional systems development model (Research Product 80-16) and Job aid manuals for phase III - Develop the instructional systems development model (Research Product 80-18). Alexandria, VA: Army Research Institute, 1980.

The overall system of which these two manuals are a part is a job aid system for the activities identified in the instructional systems development model (Branson, et al., 1975). Job aids are available for each of the five phases of the ISD model--Analyze, Design, Develop, Implement, and Control. Each job aid is composed of a descriptive authoring flowchart and a job aid manual. These volumes contain an introduction to the use of the job aid and job aid manuals, as well as the job aid manuals themselves. The manuals reproduce the flowcharts in reduced size, a segment at a time, giving necessary explanations and examples of the forms to be used.

The volume for Phase II - Design covers ISD Blocks II.1 through II.4. The volume for Phase III - Develop covers Blocks III.1 through III.5. The descriptive authoring flowcharts for each phase are available in companion documents.

Scriven, M. Evaluation bias and its control. Paper #4 in occasional paper series. Washington, D. C.: National Science Foundation, 1975. (ED 164 593)

Selected aspects of the problem of obtaining unbiased program or product evaluation are discussed. An evaluator who is a member of the project staff will have difficulty producing an evaluation which is credible and valid. Project monitors will also have a problem since they are often required to assume the conflicting roles of external evaluator and project advocate. Therefore, no unit should rely entirely on a given subunit for evaluative feedback about that same subunit. Evaluative feedback systems require renewal or replacement to prevent deterioration of their independence. Evaluators should arrange for replication of their own work by independent evaluators. Four further approaches for reducing bias in evaluation include: a. standardizing the qualitative aspects of evaluation procedures by using a checklist; b. upgrading evaluator training procedures; c. reducing sources of bias external to the evaluator; and d. comparing the project, programs, or products with alternatives.

Seidel, R.J. Transactional evaluation: Assessing human interactions during program development (HumRRO PP-8-78). Alexandria, VA: Human Resources Research Organization, 1978. (ED 159 579)

Evaluation in educational research and development programs tends to focus on the object of the innovation. Transactional evaluation focuses upon an area which is missing from these evaluative approaches—the interpersonal effects of the perceptions of project team members and the people in the environment surrounding the implementation or experimentation. The steps followed in a complete and comprehensive transactional evaluation are outlined, and examples are provided from the fields of education as well as clinical and training related settings.

Shriver, E.L. Fully proceduralized job performance aids: Guidance for performing behavioral analyses of tasks (AFHRL-TR-75-38). Brooks Air Force Base, TX: Air Force Human Resources Laboratory, 1975. (AD A015 059)

The initial tryout of fully proceduralized job performance aids (FPJPA) for the UH-1H helicopter indicated that although they met all of the format requirements for FPJPA, they did not produce the expected level of task performance when used by novice and apprentice Air Force maintenance personnel. The author hypothesized that the FPJPA did not contain all of the cues and directions necessary for the novice and apprentice personnel. In this report he describes a method for

identifying such cues and responses during a "hands on" tryout of the initially produced task steps. He calls this method the behavioral analyses of tasks (BAT). The application of this BAT to many tasks produced an "unfolding" effect from pictorial to pictorial. It also identified many important but unplanned cues in the troubleshooting routines. Its application to the eleven UH-1H tasks used for the evaluation raised the performance level of both novice and apprentice personnel. FPJPA of reasonable effectiveness will probably be developed with less rigorous "hands on" analyses of tasks than the BAT proposed in this report; provided, the FPJPA so developed are followed by a "cut and try" process of improvement. The accomplishment of a BAT requires highly skilled and tedious work on the part of each task analyst and its use will probably be viewed by some as too expensive. But the author's experience indicates that its timely use in the FPJPA development cycle will be necessary for the consistent production of a quality product at a minimum cost.

Siegel, A.I., Bergman, B.A., Federman, P., Sellman, W.S. Some techniques for the evaluation of technical training courses and students (AFHRL TR-72-15). Wayne, PA: Applied Psychological Services, 1972. (AD 753 094)

The handbook presents methods, concepts, and considerations to be held in mind in planning and implementing a student measurement or training evaluation program. Techniques are presented, procedures are discussed, and computational examples are included. The text places principal emphasis on basic techniques, but certain more advanced approaches are also considered.

Siegel, A.I., Musetti, L.L., Federman, P.J., Pfeiffer, M.G., & Wiesen, J.P. Criterion referenced testing: Review, evaluation, and extension (AFHRL TR-78-71). Wayne, PA: Applied Psychological Services, 1979. (AD A074 539)

The literature relative to criterion referenced test development is reviewed. Rater error in criterion referenced performance evaluation is discussed, and a statistical model for reducing such bias in Air Force applications is presented and experimentally evaluated. The results suggest the utility and applicability of the method in Air Force applications. Needed research into criterion referenced testing in the Air Force is described. The results of a field study into criterion referenced testing in Air Force technical training courses are presented, and the implications of the results for Air Force technical training are given.

Smith, R.G., Jr. Controlling the quality of training (DA-PROJ-2J024701A712-01-TR-65-6). Alexandria, VA: George Washington University, Human Resources Research Office, 1965. (AD 618 737)

The need for a quality control system in a military training program and the methods of establishing such a unit are described and evaluated in this report. The purpose of quality control is to insure a satisfactory standard of competence among the students who graduate, to maintain this quality by a continuous monitoring process, and to improve training where it is found to be deficient. In order to function successfully, a quality control system should constitute a separate unit which is independent of, but cooperates with, the instructional departments. Proficiency testing is viewed as the chief means of measuring the success of the training program, with emphasis on a uniform standard and a consistent method of preparation, administration, and scoring of tests.

Steiner, R. New use for assessment centers--training evaluation. Personnel Journal, 1975, 54, 236-237, 248. (ED 113 844)

The assessment center can be a mechanism for providing a highly sophisticated evaluation of the training effort. The article describes how the training manager can successfully incorporate the assessment center concept into an overall training evaluation strategy.

Swezey, R.W., & Pearlstein, R.B. Guidebook for developing criterion-referenced tests. Alexandria, VA: U. S. Army Research Institute for the Behavioral and Social Sciences, 1975. (AD A014 987)

This manual outlines the rationale for using the criterion-referenced test (CRT) approach and suggests specific guidelines for test developers to use in constructing test items. Methods for assessing the adequacy of a CRT are also provided.

The manual includes chapters on the uses of CRTs, assessment of objectives, development of a test plan, constructing an item pool, selecting the final items, administering and scoring CRTs, and assessing reliability and validity.

Tennyson, R.D. Evaluation technology in instructional development. Journal of Instructional Development, 1978, 2, 19-26.

Since Glaser's original model for instructional development, a number of others have appeared. Even the newest of these, however, offer the same general characteristics. With this in mind, the author presents

another model of instructional development, but rather than describing a well defined set of procedures, this model identifies the principles and theories associated with the process of instructional development. The assumption made here is, that with each application of these principles and theories, a unique sequencing of the development activities will be produced.

Another difference between this model and earlier models is that evaluation is included in every phase of development. In most instructional development models, evaluation has been relegated to the last step in the process--this has usually occurred because of the limitations of flowchart techniques to describe the complexity of both development and evaluation.

This model has four phases: assessment, design, production, and implementation. Within each of these phases are two main activities, development and evaluation. Additionally, the model includes reference to products associated with each phase. As a description of the model, development and evaluation activities are reviewed and suggested for use. It is assumed, however, that each developer will define his own specific strategy of development based upon an analysis of his particular situation.

Texas Southern University, Urban Resources Center. A guide to program evaluation. Volume I. Washington, D. C.: Department of Housing and Urban Development, 1974. (ED 130 404)

This booklet identifies basic techniques and methodologies for evaluating training and development programs. It also examines a selected number of methodologies that appear to be applicable to evaluating the results of program performance and achievement in many social and economic programs. Section 1 provides general information on evaluation and reviews some basic assumptions about evaluative research. Section 2 gives a brief explanation of the process of evaluation. Section 3 summarizes several selected program evaluation models, with emphasis on comparing the models and simplifying evaluation designs. Section 4 reviews other approaches to evaluation, and the appendix contains an extensive bibliography, as well as a sample evaluation form.

U. S. Civil Service Commission, Bureau of Training, Training Leadership Division. A process for the evaluation of training (U. S. Civil Service Commission Publication No. TLP-316). Washington, D. C.: U. S. Government Printing Office, 1978. (LD 004 043)

This evaluation process consists of three phases: focus, plan, and implement. Focus comprises the work which establishes the general scope of the evaluation effort. It is here that the evaluator studies

the training program to be evaluated and works with the client to establish the extent of the evaluation.

Planning produces the blueprint for conducting the project. The plan specifies the data to be collected; the sources from which they will be obtained; the times at which measurements will be taken; the ways in which the data will be gathered; and the schedule for accomplishing all necessary tasks.

During the implementation phase the evaluation plan is put into action. Activities include collecting, tabulating, and analyzing data; formulating conclusions and recommendations; and organizing and presenting the results.

U. S. Navy. Tests, measurement of student achievement. Data Item Description number DI-H-2033A, 1 Oct 1976.

This data item description gives requirements for the production of test items and tests prepared under U. S. Navy contracts dealing with system/equipment training. Rules for the construction of multiple choice, true-false, completion, cluster true-false, matching, labeling, and performance test items are given.

Wagner, H. & Seidel, R.J. Program evaluation. In H. F. O'Neil, Jr. (Ed.), Learning strategies (Ch. 8). New York: Academic Press, 1978.

The chapter deals with the evaluation of a group of learning strategies projects recently supported by the Defense Advanced Research Projects Agency (DARPA). The theme of this chapter is the application of an evaluation model that focuses on the perceptions of a program's participants (transactional evaluation) to formative evaluation of the DARPA research program. First, the authors give an overview of models of evaluation extant in education and training. They distinguish among these approaches as they apply to the formative or developmental process of a project and as they are appropriate to summative or final evaluation of completed projects or programs. In the second section, transactional evaluation is introduced as a means for dealing with a significant, though neglected, area which should be taken into account during a formative evaluation. Transactional evaluation draws on the perceptions of the project participants as indices of clarity of goals and project status during the formative stages of the project. Its importance comes from its emphasis on making explicit the relationships, roles, problems, and possible solutions as perceived by developers and potential users of a project's products. Last, the specific application of transactional evaluation to the DARPA learning strategies research program is discussed.

Webster, W.J. The evaluation of instructional materials. Washington, D. C.: Association for Educational Communications and Technology, 1976.

The evaluation of instructional materials is an extremely important yet often overlooked component of the total instructional process. Often the term evaluation is operationally defined as "checklist" by many curriculum specialists who perform a weak form of input evaluation on instructional materials, using some variation of a survey form or checklist. In presenting alternatives to inadequate evaluation, this paper attempts to accomplish three objectives:

a. To synthesize, through a selective review of the literature, a brief description of the state of the art of evaluation.

b. To present a working model demonstrating the functions of various forms of evaluation in assessing the relative merits of instructional materials.

c. To provide an annotated bibliography of sources for readers seeking further information on evaluation.

Wentling, T.L., & Lawson, T.E. Evaluating occupational education and training programs. Boston: Allyn and Bacon, Inc., Longwood Division, 1975.

Designed to serve as a handbook and guide, this comprehensive book addresses itself to educational evaluation for teachers and administrators of occupational education in public elementary/secondary/postsecondary programs and for administrators and personnel connected with private instructional programs in schools and industrial programs. However, the methodology is also adaptable to the evaluation of other academic programs. An introductory chapter briefly summarizes the history of evaluation and presents two current, widely accepted definitions of evaluation, the decision-oriented definition (Phi Delta Kappa Commission on Evaluation) and the evaluator judgement definition (Worthen and Sanders). The remaining nine chapters of the book are directed to improving the utilization of evaluation procedures, with the end result of improving decision-making and, ultimately, improving programs. Chapters 3-8 provide specific evaluation procedures: student assessment, student follow-up, employer survey, consultative team evaluation, personnel evaluation, and cost analysis. Individual chapters provide practical, class-tested evaluation activities and numerous example forms and instruments to aid in evaluation; extensive bibliographic references are included at the end of each chapter. The concluding chapter presents a general overview of how changes occur within educational programs and how evaluation results can be used to bring about change and improved programs.

Whitmore, P.G., & Fry, J.P. Soft skills: Definition, behavioral model analysis, training procedures (HumRRO PP-3-74). Arlington, VA: U. S. Army Research Institute for the Behavioral and Social Sciences, 1974. (ED 158 043)

In a report on leadership research for the U. S. Army, three papers dealing with soft skills analysis and training are presented. "What Are Soft Skills?" describes a questionnaire designed to clarify the terms "hard" and "soft" skills. Soft skills are defined as important job-related skills that involve little or no interaction with machines and whose application on the job is quite generalized. "The Behavioral Model as a Tool for Analyzing 'Soft Skills'" discusses leadership and motivation job functions in terms of principles of behavior modification and describes development of a behavioral model of the different levels of an organization. "Procedures for Implementing Soft-Skill Training in CONARC Schools" describes the instructional approach based on a tested problem-solving framework. Small groups and student-centered learning were cited as important factors in the instructional approach.

Worthen, B.R., & Sanders, J.R. Educational evaluation: Theory and practice. Belmont, CA: Wadsworth Publishing Company, 1973.

This book is a synthesis of the thinking of many leading evaluation practitioners and theoreticians. It pulls together in one volume the best of the emerging literature on educational evaluation, much of which could be found only in fugitive documents, identifies serious gaps in the literature, and provides content to fill those gaps. The result is a book which includes both the most promising conceptual frameworks proposed for educational evaluations and practical considerations in conducting such evaluations.

The book considers the state-of-the-art in educational evaluation, many frameworks and considerations in planning evaluation studies, and the future of evaluation.

Wulfeck, W.H., II, Ellis, J.A., Richards, R.E., Wood, N.D., & Merrill, M.D. The instructional quality inventory: I. Introduction and overview (NPRDC Special Report 79-3). San Diego, CA: Navy Personnel Research and Development Center, 1978 .

Also:

Ellis, J.A., Wulfeck, W.H., II, & Fredericks, P.S. The instructional quality inventory: I. User's manual (NPRDC Special Report 79-24). San Diego, CA: Navy Personnel Research and Development Center, 1979.

Fredericks, P.S. The instructional quality inventory: III. Training workbook (NPRDC Special Report 80-25). San Diego, CA: Navy Personnel Research and Development Center, 1980.

Ellis, J.A., Wulfeck, W.H., II. The instructional quality inventory: IV. Job performance aid (NPRDC Special Report 79-5). San Diego, CA: Navy Personnel Research and Development Center, 1978.

Instructional systems development, a systematic method for developing military instruction, is used by the military services to develop or revise a large portion of training courses. The Instructional Quality Inventory was developed to provide quality control/evaluation procedures for ISD.

The current IQI procedures were designed to parallel and supplement the ISD process, and are based on a system for classifying objectives, test items, and instructional presentations (the three main products of instructional development). Classification is determined according to a. what the student is required to do with the information he learns, and b. what type of information the student is learning. The IQI procedures include the following:

a. Since all ISD steps depend on careful specification of learning objectives, the first IQI procedure is to assure the adequacy of objectives. This is done by classifying each objective, and judging whether or not it accurately reflects the intended student performance after training.

b. The next step is to ensure that tests accurately measure progress toward the objectives. This is done by assessing consistency between test items and their associated objectives. Essentially, each test item must be classified in the same way as its objective. After test items and objectives are consistent, the adequacy of the test items is assessed.

c. The final step is to insure that the instructional presentation is a. consistent with the objectives and tests, and b. adequate according to psychological principles of learning.

The IQI consists of the four volumes referenced above. Volume I, Introduction and Overview, is designed to acquaint managers of instructional development efforts, evaluators of instruction, contract monitors, and others with the IQI process. Volume II, User's Manual, provides a complete description of all IQI procedures, and Volume III, Training Workbook, gives practice and feedback on IQI procedures. Volume IV, Job Performance Aid, contains brief versions of the IQI procedures.

Team Training and Evaluation

Baldwin, R.D., et al. Aircraft recognition performance of crew chiefs with or without forward observers (HumRRO TR-70-12). Alexandria, VA: Human Resources Research Organization, 1970. (AD 714 213)

A test of aircraft recognition accuracy and decision speed compared the performance of single observers and four-man crews. The test used miniaturized simulations of aircraft which were moved at scaled speeds, altitudes, and distances. The validity of the simulation was evaluated and judged by comparing the results of the test with results obtained from a previous full-scale test. Comparison of single observers with crews revealed that about 50% of the observers performed more effectively when alone than with their crew, in terms of both accuracy and decision speed. The remaining observers performed either equally well, or more effectively, when with a crew than when alone. These two groups of observers were found to prefer different communication sequences. The more effective crew observers tended to be less dependent upon other crewmen's judgements than the less effective crew observers.

Biel, W.C., Harman, H.H., & Sheldon, M.S. Exercising teams in military systems through the use of simulation (SP-1739). Santa Monica, CA: System Development Corporation, 1964. (AD 611 125)

Part I, Planning for Team Training in the System Development Process, by W. C. Biel, explains why analyses and decisions about training must be made early in the design of an operational system so that the types of training around which Parts II and III are focused can be well designed and ready when needed. Part II, Designing and Implementing the System Model, by H. H. Harman, details the design of team-training programs and reports that this design follows analysis and planning. Part III, The Evaluation of Training in a Simulated Environment, by M. S. Sheldon, discusses problems associated with the evaluation of team training or system exercising.

Briggs, G.E., & Johnston, W.A. Team training (NAVTRADEVCEN Report No. 1327-4). Orlando, FL: Naval Training Device Center, 1967.

This is the final report on a four-year program of laboratory research on team training in a Combat Information Center (CIC) context. The research literature on team training is reviewed, and a set of conclusions is drawn with regard to team performance as a function of task, training, and communication variables. In addition, the implications from this research are presented with regard to a specific team training device, the 15F5 device, which is used to teach tactical skills in the context of an airborne tactical data center.

Finally, the appendixes contain full descriptions of three laboratory studies not reported previously in the literature.

Briggs, G.E., & Naylor, J.C. Experiments on team training in a CIC-type task environment. Orlando, FL: Naval Training Device Center, 1964. (AD 608 309)

Three separate but related laboratory experiments were performed with three-man teams in a simulated radar-control interception task. Experiment I investigated the influence of a replacement of one team member with a new operator, the latter having either more or less on-the-job experience than the man replaced. Also investigated was the influence of task organization and task complexity. In Experiment II the influence of training task fidelity, training task organization, and transfer task organization was examined. Finally, Experiment III examined the influence of different amounts of experience on two kinds of training task organization and of transfer task organization. Replacement effects were significant but of short duration, but transfer task organization effects were of longer duration with performance on an independent task organization superior to that on an interaction version except when preceded by individual training and/or training specifically on communication procedures.

Brokenburr, J.L. Learning curves and their applicability to unit training levels in operational testing (Master's thesis). Atlanta, GA: Georgia Institute of Technology, 1978. (AD A086 174)

This research addresses the problem of determining the existence of a representative group/crew learning curve (or set of curves) and the development of a mathematical description of this curve applicable to training levels in operational testing. Emphasis is placed on the analysis of data from actual operational test reports. An iterative procedure is developed to analyze sample data using regression techniques to screen data for suitability and to fit nonlinear learning models. A representative learning curve for the data analyzed is selected by comparing the sum of squares regression and the lack of fit ratio for each model. This comparison shows that several models appeared to provide an adequate fit to the data analyzed. One of these adequately fits the empirical data analyzed and can be used as a representative group/crew learning model for this data.

Ciley, C.D., Jr., & Long, G.E. Development of unit training and evaluation techniques for combat-ready helicopter pilots: Task 2. Assessment of ARTEP and ATM training objectives and requirements for maintaining operational readiness (FTR 01-78). Alexandria, VA: U. S. Army Research Institute for the Behavioral and Social Sciences, 1979. (AD A069 224)

The recently published Army training and evaluation programs and aircrew training manuals represent a new concept of Army aviation unit training. Commanders are now responsible for determining the training requirements of their individual units and for developing and implementing programs to meet those requirements. The ARTEP and ATMs were designed to assist the unit commanders in carrying out that responsibility. This report presents the results of a brief review of the utility of these documents in the field and the extent to which their content adequately represents the training objectives and requirements for maintaining combat readiness. The research concludes that the documents have been well received and are being utilized effectively by field commanders; that they contain a valid, though not entirely complete, reflection of combat-readiness training objectives and requirements: but, that the required recordkeeping is burdensome and there is a need for a more effective feedback system between its users and its developers.

Collins, J.J. A study of potential contributions of small group behavior research to team training technology development. Alexandria, VA: Essex Corporation, 1977. (AD A043 911)

A review of the small group behavior research literature revealed numerout potential contributions to team training technology development. Theories, methods and techniques, and findings on substantive variables in group interaction, group performance and productivity, group growth and development, and group motives and goals are emphasized. Research program recommendations are presented. An annotated bibliography is included.

Griffith, D. An overview of the one station unit training (OSUT) attitudinal surveys (ARI Res Problems Rev 78-4). Alexandria, VA: U. S. Army Research Institute for the Behavioral and Social Sciences, 1978. (AD A076 710)

TRADOC directed that a test be conducted to determine the feasibility of replacing the current 16/17-week basic combat training (BCT)/advanced individual training (AIT) program with a 12- to 15-week one station unit training (OSUT) concept. The TRADOC test plan included a requirement for an attitude survey by the Army Research Institute (ARI) to assess trainee attitudes toward OSUT and to compare the attitudes of trainees graduating from the current BCT/AIT program with the attitudes of trainees graduating from the OSUT program. This report provides a summary of the OSUT surveys conducted at Fort Leonard Wood (MOS 12B), Fort Gordon (MOS 36C), Fort Knox (MOS 11D/E), Fort Sill (MOS 13B), and Fort Bliss (MOS 16P). For the purpose of analysis, each of 53 questionnaire items was classified into one of the following categories: background, training intensity, ancillary training, morale, reenlistment, and OSUT opinion. Two comparisons were of primary interest: BCT versus AIT and AIT versus OSUT. Items were analyzed individually. Chi-square tests were used to determine if differences in the pattern of responding between groups were statistically significant.

Hall, E.R., & Rizzo, W.A. An assessment of U. S. Navy tactical team training. Final report (TAEG R-18). Orlando, FL: Training Analysis and Evaluation Group, 1975. (ED 107 303)

A study was conducted to compile resource information for planning regarding Navy tactical team training. The specific objectives were to describe the current status of team training within the fleet; review and evaluate the findings in the technical literature regarding team training; and develop and recommend potential soulutions to team training problems. Information required for the study was gathered from two principal sources: Navy units where team training is conducted and the technical literature pertaining to team training.

Hammell, T.J., & Mara, T.D. Application of decision making and team training research to operational training. A translative technique. Orlando, FL: Naval Training Device Center, 1970. (AD 871 984)

A technique was developed to translate findings of laboratory decision making research into a form applicable to the operational anti-submarine warfare/anti-aircraft warfare training environment. This translative technique is composed of two categorization schemas - a decision skill taxonomy and a behavioral deficiency taxonomy - through which the experimental tasks studied and resultant research findings are translated. Applicability of the translated research findings to operational systems is demonstrated by an analysis of submarine fire control data from training device and real-world exercises. Data and associated information from the operational analysis are classified and unpublished.

Havron, M.D., & Wanschura, R.G. Improved Army training and evaluation program (ARTEP) methods for unit evaluation. Volume VII. Executive summary (HSR RR-79/4-GE; ARI TR-79-A25). Alexandria, VA: U. S. Army Research Institute for the Behavioral and Social Sciences, 1979. (AD A076 957)

This report summarizes a two phase, 28-month contract research effort concerning the ARTEP for tank/mechanized infantry units. Products and processes of the research are described. Products are represented by seven report documents, including: ARTEP implementation problem diagnosis and issue identification, analysis of issues and concepts for solution, exercise planning guidance, evaluator/controller training, analysis of alternative training settings in the tank/mechanized infantry battalion training environment, and integration of engagement simulation training methods into unit evaluation focused field exercises.

Jacobs, T.O., Ward, J.S., Powers, T.R., George, C.E., & McFann, H.H. Individual and small unit training for combat operations (Professional Paper 21-67). Alexandria, VA: George Washington University, Human Resources Research Office, 1967. (AD 653 845)

Contents: Training for modern combat operations; a case study of the development of an individual combat training program; the foundations for leader training; training for coordination within rifle squads; and individualization of instruction.

King, W.C. Team training: A review of selected literature. Norfolk, VA: Calspan Corporation, Advanced Technology Center, Instructional Systems Office, 1980.

The scientific study of team training problems for the United States military began in earnest in the 1950's, with programs at System Development Corporation, American Institutes for Research, the Air Force Personnel and Training Research Center, and at Ohio State under Naval Training Device Center contracts (Briggs & Johnston, 1967). This review touches on some of this early research, to the extent that it is still relevant, and traces major themes in team training research and development up to the present. The basic issues addressed are:

a. The goals and advantages of team training.

b. Major problems and difficulties in team training.

c. Possible solutions to the above problems, suggested by the literature.

d. Recent developments in team training.

e. Suggestions for team training research.

Klaus, D.J., & Glaser, R. Increasing team proficiency through training: 5. Team learning as a function of member learning characteristics and practice conditions (AIR E1-4/65-TK). Pittsburgh, PA: American Institutes for Research, Team Training Laboratory, 1965. (AD 471 469)

This study investigated the effects of variations in team member characteristics and team practice conditions on the rate at which a team response is acquired and extinguished. The variables investigated included: a. the average response proficiency attained by individual team members at the time the team was formed, b. the rate at which this level of member proficiency was attained during individual practice, c. the degree of homogeneity in proficiency among the members comprising a given team, and d. the extent of delay between the completion of individual learning and the initiation of team training. Each of the 28

teams studied was organized in a modified series arrangement so that all three members had to be correct for a team reinforcement to occur. Of the variables studied, only the proficiency level of the members at the initiation of team training was a determinant of the rate of team acquisition or team extinction. Supplemental analyses, however, revealed several differences in the course of learning, aside from rate, which were attributable to the other variables.

Klaus, D.J., & Glaser, R. Increasing team proficiency through training (AIR E1-6/68-FR). Pittsburgh, PA: American Institutes for Research, Team Training Laboratory, 1968. (AD 669 688)

The report summarizes the results of research at the Team Training Laboratory from December 1960 until August 1967. During this time, seven technical reports were issued by the laboratory. This summary report briefly describes each of these seven studies and reviews their purpose and major results. The final section of this report identifies some practical implications of this research and relates the underlying concepts to the broader context of social behavior.

Klaus, D.J., & Glaser, R. Reinforcement determinants of team proficiency. Organizational Behavior and Human Performance, 1970, 5, 33-67.

The proficiency of working teams having well-defined structures and member assignments can be considered as a function of the occurrence of reinforcement for the group as a whole following each team response. Findings from a series of seven studies suggest that increments and decrements in team performance are predictable from a knowledge of reinforcement contingencies and team structures and may be attributed to the differential effects of group reinforcement on individual team members. Both "series" teams, requiring specified contributions from all members, and "parallel" teams, those containing redundant members, were studied. The effects of characteristic entering performance, supplementary feedback during team training and the simulation of team conditions during the training of individuals also was investigated. Some implications of the research and the underlying model are identified with respect to the broader context of social behavior.

Klaus, D.J., Grant, L.D., & Glaser, R. Increasing team proficiency through training: 6. Supervisory furnished reinforcement in team training (AIR E1-5/65-TR). Pittsburg, PA: American Institutes for Research, Team Training Laboratory, 1965. (AD 471 470)

As a consequence of the reduced ratio of reinforcement in going from individual to subsequent team training, initial levels of team

proficiency are far lower than predicted on the basis of the proficiencies of individual members. To overcome this decrement, experimental teams were provided with both individual and team reinforcement during the initial stages of team training. Results were obtained from 24 laboratory teams at varying levels of proficiency. The combined use of team and individual reinforcement produced team proficiency more rapidly than when team training was given without individual reinforcement. The experimental teams required almost as many trials to achieve the team proficiency criterion under team reinforcement alone after having once achieved it with combined team and individual reinforcement as did the control teams. The main conclusion is that the use of supervisory-furnished individual reinforcement on a regular basis during team training is satisfactory only if also provided later in the performance situation.

Kribs, D., et al. Computerized collective training for teams (ARI TR-77-A4). Arlington, VA: U. S. Army Research Institute for the Behavioral and Social Sciences, 1977. (ED 140 854)

A review and evaluation was conducted of state-of-the-art findings and instructional theory directly applicable to the problem of developing instructional strategies for computer-assisted team training. Two major conclusions were drawn from the review and evaluation. The first is that a conceptual framework for a general purpose set of instruction strategies for team training does not exist. The second conclusion is that an instructional systems development apporach must be developed for team training. The paper also describes the initial development for deriving team training instructional strategies. Three major elements were identified: a. team task dimensions and team training objectives; b. learner characteristics and strategies; and c. characteristics of the training delivery system used to implement the strategies.

Larson, O.A. Survey of unit performance effectiveness measures (NPRDC TR-74-11). San Diego, CA: Navy Personnel Research and Development Center, 1974. (AD 774 919)

Improved measures of performance effectiveness are required by the Marine Corps for its combat unit training program in order to ensure the maintenance of appropriate levels of unit readiness in accordance with its assigned mission. A survey to determine the state-of-the-art of performance assessment systems and methodologies was conducted as an initial research phase in support of this requirement. A two-fold effort was made to review the research literature in performance evaluation, decision making, and unit training, and to gather first-hand information about existing performance assessment systems. A broad informational survey provides a number of alternative theoretical and practical methodologies which may serve as feasible approaches in ensuing research. The final 20 pages of the report are devoted to an annotated bibliography.

Long, G.E., Ciley, C.D., Jr., Hockenberger, R.L., & Garlichs, E.A. Development of unit training and evaluation techniques for combat-ready helicopter pilots: Task 1. Development of an instruction program for individual and unit training with combat-ready pilots (FTR 05-78). Alexandria, VA: U. S. Army Research Institute for the Behavioral and Social Sciences, 1979. (AD A069 242)

The attainment and maintenance of combat readiness must be a primary responsibility of every field unit commander. To carry out that responsibility, commanders of units with combat missions need training and evaluation techniques that will train pilots who already know how to fly to accomplish the requirements of a specific mission in a battlefield environment as part of the combined arms team. In order to provide the required techniques and procedures noted above, it is first necessary to derive an effective and efficient approach to their development. That approach should result in the availability of techniques and procedures that will facilitate the attainment of the highest level of combat readiness in the largest number of operational units in the shortest amount of time. This report describes the research effort directed at the derivation of an approach to the development of training and evaluation techniques and procedures for combat-readiness training that will meet the above noted requirements. It also describes the effort directed at the partial development of two training modules consistent with that approach.

Morgan, B.B., Jr., Coates, G.D., Alluisi, E.A., & Kirby, R.H. The team-training load as a parameter of effectiveness for collective training in units (ITR 78-14). Norfolk, VA: Old Dominion University, Performance Assessment Lab, 1978. (AD A063 165)

This report summarizes the results of two series of studies of team training conducted during the summer of 1977. In each of 10 studies, 5 subjects worked together as a team for 8 hours per day over 6 consecutive days; during their first 48 hours of work, each team was trained to perform the 6 tasks that constitute the synthetic work presented with the Multiple-Task Performance Battery. The 10 teams consisted of different combinations of the total of 20 undergraduate male volunteer subjects to provide team-training loads (percentages of untrained team personnel) ranging from 0 to 100 percent in 20 percent steps. The data of the 10 studies were combined to permit analysis of the effects of team-training loads ranging from 0 to 100 percent in 10 percent steps, and the effects of team-training load on training and performance effectiveness were thereby assessed.

Northwest Regional Educational Laboratory. Evaluation design for social conflict and negotiative problem solving. Washington, D. C.: National Institute of Education, 1976. (ED 127 343)

Social Conflict and Negotiative Problem Solving is an instructional system developed by the Improving Teaching Competencies Program of Northwest Regional Educational Laboratory. This report presents a plan of evaluation activies for the interim milestone period in the development of the instructional system. Social Conflict and Negotiative Problem Solving is designed for teachers, administrators, and others to increase their ability to recognize and handle conflict due to differences of values and self-interest. It is intended to be a relatively structured, experience-based workshop designed to provide a variety of opportunities to explore situations of social conflict. The training is designed to provide conceptual awareness and experiential training in the following areas: social conflict, power, assertiveness, self-interests, interpersonal communication skills in conflict situations, and "negotiative" problem solving skills.

Payne, W.H., & Braunstein, D.N. Suitability of a simple task for the study of team training problems (SRM 65-5). San Diego, CA: Naval Personnel Research Activity, 1965. (AD 466 192)

Four teams of four subjects were given a signal detection task under three organizing conditions. Stimulus materials were rigidly controlled, and order of conditions was counterbalanced. No significant differences were found in number of signals detected. Experiments using similarly controlled stimulus materials, but involving more complex tasks and organizing conditions, are suggested in order to study the effect of team organizing conditions on performance.

Popelka, B.A., & Knerr, C.M. Team training applications of voice processing technology. Springfield, VA: Litton Mellonics Systems Development Division, 1980. (AD A085 999)

Automated speech technology and intelligent computer assisted instruction offer unique solutions to problems of training teams in communication and coordination skills. At this point in the emergence of automated speech technology, scientists have only begun to explore its training uses. The application of automated speech technology entails adaptive training, or intelligent computer assisted instruction techniques in which the computer acts like a human tutor. This report reviews the goals and accomplishments of automated speech processing and its application to training, especially military team training.

Reeves, J.M. & Michael, W.B. Application of the Stufflebeam educational decision-making model to the evaluation of a dental team training program involving the use of paraprofessionals. Paper presented at the Northeastern Educational Research Association Convocation, Ellenville, NY, 1973. (ED 094 004)

The results from application of Stufflebeam's comprehensive decision-making methodology--the context-input-process-product evaluation model--to the evaluation of a dental team training program with expanded functions of auxiliary personnel (paraprofessionals) at a school of dentistry are described. This training program was aimed at enhancing the capabilities of the dental profession to serve a large public at lower cost and with greater efficiency without a reduction of quality in service. For this innovative training program, it appeared essential to delineate, obtain, and provide useful information for: a. planning decisions regarding appropriate change-oriented objectives based on a rationale of needs (facilitated by context evaluation); b. structuring decisions concerning a choice of alternative designs, strategies, and procedures for conducting the program (served by input evaluation); c. implementing decisions pertaining to carrying out and monitoring the program design and strategy (directed by process evaluation); and d. recycling decisions concerning whether to continue, terminate, modify, or refocus the change activity (aided by product evaluation).

Short, J.G., Cotton, T., & Klaus, D.J. Increasing team proficiency through training: 7. The simulation of team environments (AIR E1-5/68-TR). Pittsburgh, PA: American Institutes for Research, Team Training Laboratory, 1968. (AD 669 687)

Three studies of simulated team environments are described in the report. Each study dealt in some way with the transition performance decrement that occurs when individuals are placed in teams. Study 1 found that this decrement was, in a large part, a function of the change in reinforcement ratio that occurred between individual and team training. Study 2 showed that reinforcement in the team environment was a function of characteristics of the team members themselves--their number and the proficiencies. An increase in the number of team members or a decrease in their proficiency produced lower reinforcement ratios, and these in turn produced larger decrements in performance and increased the time required for teams to reach high levels of proficiency. Study 3 showed that the transition performance decrement could be reduced by certain training strategies notably those involving a simulation of the team environment during individual training.

Shriver, E.L., Mathers, B.L., Griffin, G.R., Word, L.E., & Root, R.T. REALTRAIN: A new method for tactical training of small units. Alexandria, VA: U. S. Army Research Institute for the Behavioral and Social Sciences, 1975. (AD A024 030)

REALTRAIN is an improved, low-cost tactical training and evaluation technique for use in Army combat unit training exercises. Realistic, two-sided, free-play tactical training employing recognized principles of learning is achieved through simulated combat engagement situations.

For Infantry, a soldier with a 6X telescope mounted on an M16 rifle attempts to identify 3-inch numbers on the helmet of his "enemy." When the number is identified, he fires a blank round and reports the "hit" to a controller who is in constant communication with controllers on the other side, and "the enemy" is withdrawn from action. An after action review, in which the participants describe and discuss their roles in the action, reinforces the lessons learned. The method is enthusiastically accepted and the learning of appropriate behavior is rapid. REALTRAIN principles have also been successfully adapted to Armor and Antiarmor units.

This report describes the development of the REALTRAIN training method, which incorporates the casualty assessment techniques into an appropriate learning environment; discusses in detail major aspects of the training method; and presents data on the effectiveness of the method, its acceptance by men in units in the field, and its utilization in Army units throughout the world. This report also discusses future research in the refinement and extension of the REALTRAIN method.

Smith, E.A. Four systems for controlling multiscreen or team training presentations. Final Report (AFHRL TR-77-83). Brooks Air Force Base, TX: Air Force Human Resources Laboratory, 1977. (ED 160 066)

The four instructional systems described consist of a. a system for controlling multiple images that can be assembled in modular fashion starting with existing equipment and systematically adding components as additional functions are required; b. a more complex system for controlling multiscreen presentations that requires a considerable initial investment of money and personnel; c. a team training configuration designed to provide orientation or theory to small teams; and d. a configuration for providing performance oriented training to teams. Discussions are limited to techniques for implementing the instructional strategy, with no attempt to present data regarding instructional effectiveness. A summary of a classroom field test and evaluation of the usability of the modular configuration is included. An operations manual for the complex multiscreen system, and a description of the course development of the team training package are appended.

Thurmond, P. Development of analysis, design, and development techniques for team ISD. Alexandria, VA: U. S. Army Research Institute for the Behavioral and Social Sciences, 1980.

The three tasks set forth for this investigation were a. analyze team operations, and identify team structure and processes relative to the development of team training, b. determine the applicability of the interservice procedures for instructional systems development (ISD) to the analysis, design, and development of team instructional materials, and c. identify appropriate procedures for the analysis, design, and development of team instructional materials.

It was found that none of the ISD procedures currently employed are adequate as guidelines for team training. Critical areas where documentation is deficient include collective front-end analysis (CFEA), identification of types of team learning, and development of collective training scenarios that incorporate efficient learning principles and represent actual mission contingencies. Recommendations for improving collective training analysis, design, and development are included.

Thurmond, P., & Kribs, H.D. Computerized collective training for teams. Final report (ARI TR-78-A1). Arlington, VA: U. S. Army Research Institute for the Behavioral and Social Sciences, 1978. (ED 162 629)

The purpose of this investigation was to empirically demonstrate and evaluate a brassboard for computerized collective training for teams (COLT2). The underlying tasks were to a. conduct a state-of-the-art assessment of instructional strategies appropriate for COLT2, b. derive a conceptual framework for COLT2 instructional strategies, c. conduct a team job/task and training analysis for COLT2 on the Army computerized artillery fire control system (TACFIRE), and d. develop TACFIRE team training scenarios for the purpose of instructional strategy assessment. The procedures included design and implementation of a team ISD model, via which sample training materials were developed. The materials were adapted to the team training version of the PLANIT CAI system. The results of the developmental aspects of the project indicate that many of the components of the team ISD approach that were designed for this effort would adequately meet the criteria for a generic team ISD model. Preparation of team learning objectives and evaluation of interactive team skills are deficient areas. PLANIT met the basic team CAI requirements. The results preliminarily indicate that there are differences in regard to what types of behavior are learned between team and individual instruction.

U. S. Army Infantry School. How to prepare and conduct military training (FM 21-6). Fort Benning, GA: U. S. Army Combat Arms Training Board, 1975. (ED 132 374)

Designed to apply to any unit regardless of strength, mission, organization, or equipment, this field manual provides trainers with methods and techniques of preparing and conducting individual and collective training. Chapter 1 discusses the purpose and scope of this publication and explains the duties and relationship between the trainer (for whom this manual was written) and the training manager. Chapter 2 discusses the purpose of training. Chapter 3 describes a three-step, backward planning process to prepare, conduct, and evaluate training of individuals to perform their duty assignment. Chapter 4 is an introduction to collective training, which prepares soldiers to perform those team or unit tasks essential to the accomplishment of a unit's operational missions.

Chapter 5 offers examples which show company level trainers how the fundamentals of collective training are used to prepare and conduct equipment-oriented collective training. Chapter 6 deals with tactical collective training. The eight appendixes include: training publications; practical exercises in writing training objectives; training techniques, aids, and devices; evaluating (inspecting) training; tactical exercises; training trainers to train; sample lesson plans; and an index and glossary of training terms.

Wagner, H., Hibbits, N., Rosenblatt, R., & Schulz, R. Team training and evaluation strategies: State-of-the-art (HumRRO-TR-77-1). Alexandria, VA: Human Resources Research Organization, 1977.

A critical review and evaluation of the literature was performed to describe existing instructional and evaluative techniques relevant to team training. Comprehensive documentation sources such as Educational Resources Information Center, Defense Documentation Center, and National Technical Information Service were searched. In addition, the social psychological areas and the industrial training field were surveyed for relevant publications. Current studies and team training practices underway within the Services were examined and described. A classification scheme was introduced in which the training situations discussed in this review were categorized as either "emergent" or "established." Established situations are those in which the tasks and required behaviors can be almost completely specified. Emergent situations permit some discretionary behaviors because all activities cannot be predicted. Additionally, "team" training was distinguished from "multi-individual" training. Although both occur in a group setting, the focus in the latter is on individual skills, whereas team training focuses on team skills (e.g., coordination). Using this classification scheme, state-of-the-art gaps in team training and evaluation were identified for needed research and development. In addition, new techniques, such as simulated two-sided engagement training, were suggested as warranting further study of their applicability to team training.

APPENDIX B

Short Forms Of Master List Questions

1.	Is the TEST FORMAT appropriate for the OBJECTIVE?	(See table in job aid.)
2.	Are there test items for the TLO or all of its critical parts/LOs? (See job aid for critical parts.)	1 = Test items for TLO or all parts/LOs 2 = No test items for TLO and for some parts/LOs 3 = No test items for TLO and for most parts/LOs
3.	Is there a test item for each critical part of each LO? (See job aid for critical parts.)	1 = Items for all parts 2 = Items for many, but not all parts 3 = Items for only a few parts or for no parts
4.	Does the TASK LEVEL of the test item match the TASK LEVEL of its OBJECTIVE?	(See table in job aid.)
5.	Does the content of the test item match the content of its OBJECTIVE ?	1 = Same 2 = Slightly different 3 = Very different
6.	Do the CONDITIONS of the test item match the CONDITIONS of its OBJECTIVE?	1 = Exact match 2 = Minor mismatch 3 = Severe mismatch
7.	Do the STANDARDS of the test item match the STANDARDS of its OBJECTIVE?	1 = Exact match 2 = Minor mismatch 3 = Severe mismatch
8.	For true-false, multiple choice, and matching items is only one answer correct?	1 = Only one answer is correct 3 = More than one answer can be correct.
9.	For short answer, fill-in, listing, and performance items are all acceptable answers in the answer key?	1 = All correct answers are in answer key. 3 = Some correct answers are <u>not</u> in answer key.
10.	Does the test item provide opportunities for common errors to be made?	1 = Yes 3 = No
11.	Is the language of the test item easy for students to understand?	1 = Easy 2 = Somewhat difficult 3 = Very difficult

12. Is the test item different from previous PRACTICE and EXAMPLES? (USE-CONCEPT, USE-RULE, or USE-PRINCIPLE only)	1 = Different 2 = Presented before, USE-UNAIDED 3 = Presented before, USE-AIDED
13. Is the answer to the test item given away by other item(s)?	1 = Answer not given away 2 = Other items give clues. 3 = Answer can be found in other item(s).
14. Is the answer to the test item dependent on answering previous item(s) correctly?	1 = Answer not dependent on other items 3 = Previous items must be correctly answered.
15. Are sketches and diagrams used in the test item easy to understand?	1 = Easy to understand 2 = Somewhat confusing 3 = Very confusing
16. Is the test item tricky or misleading?	1 = Not misleading 2 = Somewhat misleading 3 = Very misleading
17. Is the test item well constructed? (See job aid for criteria list for the test format used.)	1 = Meets all criteria 2 = Deficient on a few criteria 3 = Deficient on several criteria
18. When performance steps are scored, does the instructor use a checklist?	1 = Fills in completely 2 = Uses as a reference or fills in partially 3 = Does not use
19. Is each correct answer position used about the same number of times? (true-false, multiple choice, or matching items only)	1 = Yes 3 = No
20. Are specific patterns of correct answer positions repeated across test items or are single positions repeated in blocks? (true-false, multiple choice, or matching items only)	1 = No patterns easily seen 3 = Patterns can be seen.
21. Are test administration directions complete?	1 = Directions are complete. 2 = Directions provided, but incomplete or unclear 3 = Directions are not provided

22.	Do instructors follow the directions when administering the test?	1 = Yes 2 = Some variations from directions 3 = Significant variations from directions
23.	Are adequate test instructions provided to the student?	1 = Yes 2 = Instructions provided, but unclear 3 = No instructions provided
24.	Does the final test integrate tasks as they are integrated in the "real world"?	1 = Yes 2 = Partially 3 = No, tasks are tested separately
25.	Are tasks and task steps tested in the same sequence as they are performed in the "real world"?	1 = Yes 2 = Slightly out of sequence 3 = Very different sequence
26.	Is the test free of external cues or help?	1 = Yes 2 = Hints given. 3 = Answers are given away.
27.	Are motivational techniques employed?	1 = Yes (Please describe them.) 3 = No
28.	Is the trainees' attitude positive?	1 = Positive 2 = Indifferent 3 = Hostile or frustrated
29.	Are course ENTRY SKILLS reviewed?	1 = Review with practice 2 = Review with no practice 3 = No review
30.	Is mastery of prerequisite skills verified prior to new instruction?	1 = Yes 3 = No
31.	Are OBJECTIVES presented to the student?	1 = Yes 3 = No
32.	Are the basic PRESENTATION COMPONENTS present?	(Rated only by TEE analyst. See guidance and tables in handbook.)
33.	Are STATEMENTS complete?	1 = Statements complete 2 = Few parts missing 3 = Many parts missing
34.	Are STATEMENTS for CONCEPTS, PROCEDURES, or RULES adequate? (See job aid criteria.)	1 = Completely adequate 2 = Some or all features omitted.

35.	Does STATEMENT HELP provide sufficient explanation?	1 = Help provides sufficient explanation. 2 = Help gives insufficient explanation. 3 = Help is confusing.
36.	Does training include instruction on the use of required job performance aids?	1 = Yes 3 = No
37.	Are EXAMPLES and NON-EXAMPLES adequate?	1 = Yes 3 = No
38.	Is EXAMPLE HELP adequate?	1 = Help provides sufficient explanation. 2 = Help gives insufficient explanation. 3 = Help is confusing.
39.	Are EXAMPLES sequenced from easy to hard? (CONCEPTS only)	1 = Yes 3 = No
40.	Are there enough EXAMPLES? (See job aid for criteria.)	1 = Yes 3 = No
41.	Are NON-EXAMPLES included ? (CONCEPTS only)	1 = Yes 3 = No
42.	Do DEMONSTRATIONS show how to correct/ avoid common errors?	1 = Yes 3 = No
43.	Are steps in a DEMONSTRATION the appropriate size? (See job aid.)	1 = Yes 2 = Step size is too small. 3 = Step size is too large.
44.	Are tasks and task steps DEMONSTRATED in the same sequence as they are performed in the real world?	1 = Yes 2 = Slightly out of sequence 3 = Very different sequence
45.	Are memory aids used? (PRACTICE REMEMBERING only)	1 = Used 3 = Not used
46.	Does each PRACTICE REMEMBERING item have the same content and format as the test item?	1 = Same 2 = Same content, different format 3 = Different content
47.	Are PRACTICE USING items sequenced from easy to hard?	1 = Yes 3 = No

48.	Do PRACTICE USING items provide opportunities for COMMON ERRORS to be made?	1 = Yes 3 = No
49.	Are PRACTICE items different from EXAMPLES? (USE-CONCEPT, USE-RULE, or USE-PRINCIPLE only)	1 = Different 3 = Presented before
50.	Does PRACTICE USING integrate tasks as they are integrated in the "real world"?	1 = Yes 2 = Partially 3 = No, tasks are practiced separately.
51.	Are JOB PERFORMANCE AIDS (JPAs) usable? (See criteria in job aid.)	1 = Easy to use 2 = Hard to use 3 = Unusable
52.	Do all students use the JOB PERFORMANCE AID?	1 = Yes 2 = Some do not use JPA. 3 = Most do not use JPA.
53.	Does the TASK LEVEL of the PRACTICE item match that of the test item(s)?	1 = Yes (2 and 3--See table in job aid.)
54.	Does the CONTENT TYPE of the PRACTICE item match that of the test item(s)?	1 = Yes 3 = No
55.	Does the FORMAT of the PRACTICE item match that of the test item(s)?	1 = Yes (2 and 3--See table in job aid.)
56.	Do the CONDITIONS of each final PRACTICE item match those of the test items(s)?	1 = Yes 2 = Slightly different 3 = Very different
57.	Do the STANDARDS of each final PRACTICE item match those of the test item(s)?	1 = Yes 2 = Slightly different 3 = Very different
58.	Is final PRACTICE free of external cues of help?	1 = Yes 2 = Hints given 3 = Answers are given away.
59.	Are there PRACTICE items for each TLO or all of its critical parts/ LOs?	1 = Practice items for the TLO or all parts/LO's 2 = No Practice items for the TLO and for some parts/LOs 3 = No Practice items for the TLO and for most parts/LOs

60.	Is there a PRACTICE item for each critical part of each LO? (See job aid for critical parts.)	1 =Practice for all parts. 2 =Practice for many, but not all parts. 3 =Practice for only a few parts or for no parts.
61.	Do all students PRACTICE?	1 = Yes 2 = Some students do not Practice. 3 = Most students do not Practice
62.	Do students who PRACTICE do so until they meet the required STANDARDS?	1 = Yes 3 = No
63.	Is FEEDBACK provided for PRACTICE?	1 = Feedback Help is given. 2 = Correct answer only is given. 3 = No feedback is given.
64.	Is FEEDBACK HELP adequate?	1 = Help gives enough explanation. 2 = Help gives insufficient explanation. 3 = Help is confusing.
65.	Is TEAM PRACTICE provided?	1 = Yes 3 = No
66.	Are TEAM PRACTICE CONDITIONS the same (or as close as possible) to those of the real task?	1 = Yes 2 = Slightly different 3 = Very different
67.	Is TEAM PRACTICE FEEDBACK provided?	1 = Feedback Help is given. 2 = Success/Failure feedback only is given. 3 = No feedback is given.
68.	Is FEEDBACK HELP for TEAM PRACTICE adequate?	1 = Help gives enough explanation 2 = Help gives insufficient explanation. 3 = Help is confusing.
69.	Are all PRESENTATION COMPONENTS separated and identified?	1 = Yes 2 = Some are not 3 = Most or all are not
70.	Is the technical quality of written or spoken material adequate? (See job aid for criteria. Make notes on specific problems.)	1 = Most criteria met 2 = Several criteria not met 3 = Few criteria met

71.	Is the wording of written or spoken material easy for the students to understand?	1 = Yes, few hard words and long sentences 2 = Some hard words and long sentences 3 = Many hard words and long sentences
72.	Is the instructor's presentation or the narration easy to listen to?	1 = Yes 2 = Dull and monotonous 3 = Hard to listen to
73.	Is the instructor's presentation or the narration supported by visuals?	1 = Completely 2 = Partially 3 = Not at all
74.	Are visuals easily understood?	1 = Yes 2 = Understandable with effort 3 = Very hard to understand
75.	Are the OBJECTIVES (TLOs and LOs) within each LESSON sequenced properly? (Prerequisites taught first.)	1 = Yes 3 = No
76.	Are the LESSONS sequenced properly within the course?	1 = Yes (Rated only by THE Analyst) 3 = No
77.	Are the media appropriate for the objectives? (See table in job aid.)	1 = Yes 3 = No (Note key words, underlined in table, on worksheet.)
78.	Can the media used provide all necessary stimuli?	1 = Yes 3 = No
79.	Are the course administration directions complete?	1 = Yes 2 = Partially incomplete 3 = Incomplete or non-existent
80.	Do course administration directions make realistic demands of students and instructors?	1 = All demands are realistic. 3 = Some demands are unrealistic. (Note what they are.)
81.	Is the instructor/trainee ratio such that all students can see, hear, and receive feedback?	1 = Yes 2 = A few students cannot see, hear, and receive feedback. 3 = Many students cannot see, hear, and receive feedback.

82.	Does the instructor follow the methods in the Instructor Guide?	1 = Yes 2 = Follows to some extent. 3 = Follows very little or not at all.
83.	Does the instructor teach all of the content in the lesson materials?	1 = Yes 2 = Much of the content 3 = Very little of the content (If 2 or 3, note what was left out.)
84.	Did the instructor limit his teaching to the content in the lesson materials?	1 = Yes 3 = No (Please note what other things he taught).
85.	Is there enough space for all of the trainees?	1 = Yes 2 = A little crowded 3 = Very cramped or some students can't fit in the space at all.
86.	Is instruction free of distractions?	1 = Yes 2 = Distractions are annoying. 3 = Distractions seriously interferes with the instruction.
87.	Is the lighting appropriate for the training situation?	1 = Yes 2 = Students have trouble reading or seeing displays & equipment. 3 = Students cannot read or see displays & equipment.
88.	Is the temperature appropriate for the training situation?	1 = Yes 2 = Temperature makes students uncomfortable. 3 = Temperature seriously interferes with learning.
89.	Is the instructor's attitude positive?	1 = Yes 3 = No
90.	Are frequent breaks provided? (5-10 minute breaks every hour)	1 = Yes 2 = Breaks too short or infrequent 3 = Breaks not provided
91.	Is the speed of presentation appropriate?	1 = Yes 2 = Too slow 3 = Too fast
92.	Does the training device/equipment used in training function properly?	1 = Yes 2 = Minor malfunctions, little change from intended task performance. 3 = Major malfunctions, substantial change from intended task performance

93. Is there anything else unusual about the lesson materials, or do any other critical incidents occur during training that would interfere with learning? (Describe each one below. Rating = 3)